Enhanced Unscented Transform Method for Probabilistic Load Flow Studies

By

Oluwabukola Abiodun OKE, B.Eng, M.Sc.

Department of Electrical and Electronic Engineering

Faculty of Engineering

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy,

May 2013

Abstract

The advent of deregulated electricity and the call for sustainable energy practices are major drivers for the continued increase of renewable energy systems within the modern day power network. Dominant among them is the wind energy system whose output is uncertain because of its dependence on the prevailing climatic conditions. This increases the level of uncertainty witnessed within the power system as such, as the penetration of renewable energy systems continue to increase, their effects cannot be trivialised.

Probabilistic load flow (PLF) is employed by power system analysts to account for the effect of uncertainty within the power network. The common technique which is based on Monte Carlo Simulation (MCS), though accurate is very time consuming and for large systems it becomes unwieldy. Alternative approaches with the advantages of the MCS method but with reduced computational burden are required. A viable alternative method should therefore require minimum computational time and burden, be able to accurately model various network uncertainties, be applicable to practical small and large systems, be able to account for the effect of dependency among network variables and possess good overall accuracy.

This thesis proposes a novel approximate approach referred to as the enhanced unscented transform method to meet the requirements of PLF. The method combines the Gaussian quadrature method and the Stieltjes procedure with dimension reduction technique in deciding estimation points while the Cholesky decomposition is incorporated to account for the effect of dependency. The performance of the proposed technique is demonstrated using modified IEEE 6, 14, and 118 test systems and a practical distribution test system all incorporating wind farms. Results obtained for numerous scenarios show a good match between the proposed method and the MCS method but with significant computational burden saving. The performance of the method is also shown to compare favourably with other existing PLF methods.

List of Publications

- O.A Oke, D.W.P. Thomas, G.M. Asher and L.R.A.X de Menezes, "Probabilistic Load Flow for Distribution Systems with Wind Production using the Unscented Transforms Method," *Innovative Smart Grid Technologies (ISGT), 2011 IEEE PES,* pp.1,7, 17-19 Jan. 2011.

- O.A. Oke, D.W.P. Thomas, and G.M. Asher, "A new probabilistic load flow method for systems with wind penetration," *PowerTech, 2011 IEEE Trondheim,* pp.1-6, 19-23 June 2011.

- O.A. Oke, D.W.P. Thomas, L.R.A.X de Menezes and C. Christopoulos, "The use of unscented transforms in modeling the statistical response of nonlinear scatterer in a reverberation chamber," *XXXth General Assembly and Scientific Symposium, 2011* URSI , pp.1-4, 13-20 Aug. 2011.

- O.A. Oke and D.W.P. Thomas, "Enhanced cumulant method for probabilistic power flow in systems with wind generation," *11th International Conference on Environment and Electrical Engineering (EEEIC), 2012,* pp.849-853, 18-25 May 2012.

- O.A. Oke and D.W.P. Thomas, "Probabilistic load flow in microgrid assessment and planning studies," *Electrical Power and Energy Conference (EPEC), 2012 IEEE,* pp.151-156, 10-12 Oct. 2012.

- O.A. Oke and D.W.P. Thomas, "Probabilistic Load Flow Considering Correlation for Power Systems with Wind Generation" 12th International Conference on Probabilistic Methods Applied to Power Systems

- O.A. Oke, D.W.P. Thomas, and G.M. Asher, "Enhanced Unscented Transform Method for Probabilistic Load Flow in Systems with Wind Penetration" Submitted to *IEEE Transaction on Power Systems*

"If I have seen further it is by standing on the shoulders of Giants"- Isaac Newton

Acknowledgement

My first gratitude goes to God for His ownership by creation and recreation and for the backing given me to successfully pursue this programme.

I am indebted to my supervisor Prof David Thomas for his guidance, patience and incessant support in ensuring the completion of this work. His effort has without doubt helped through the dark periods of this research work. Special thanks to Dr Okan Ozgonenel of the Onodokuz mayis University, Samsun Turkey for providing the measured wind data and the distribution system network. I am also thankful to Dr Arthur Williams for his continuous interest in my progress and support though my stay in Nottingham.

My unreserved gratitude goes to the University of Nottingham and the Schlumberger Foundation for sponsoring this research work.

I appreciate Prof E. Owolabi, and my other tutors for their continuous interest and encouragement from my undergraduate days. I also thank my group of friends "The Elected Ones" for their encouragement and believe in me. I cannot but mention the contributions of my friends in Nottingham to my having an enjoyable stay all through my study. I am particularly indebted to Pastor Paul Akinwamide and the entire members of DLBC, Nottingham for the care and communion I found in them.

I deeply cherish my late dad for his great sacrifice, unfeigned love and care till death. I specially thank my mum for the support, care, love and instilling strong virtues in me through the years. I am eternally grateful to my wonderful siblings and in-laws for the push, unwavering support and being true role models in my quest for excellence. Finally, I am thankful to, Emmanuel *ihotu* for the support, love and the sunshine he has brought.

Content

Chapter 4

Probabilistic Load Flow Methods

Chapter 5

The Conventional Unscented Transform Method

Chapter 6

Gaussian Quadrature for Unscented Transform Method

List of Figures

List of Tables

List of Abbreviations

BDR	Bivariate Dimension Reduction
BFS	Backward Forward Sweep
CDF	Cumulative Distribution Function
CM	Cumulant Method
cUT	Conventional Unscented Transform
CV	Coefficient of Variation
DFIG	Doubly Fed Induction Generator
EU	European Union
FDLF	Fast Decoupled Load Flow
FFRPF	Fast and Flexible Radial Power Flow
FFT	Fast Fourier Transform
FOR	Forced Outage Rate
FOSMM	First Order Second Moment Method
gUT	Gaussian Quadrature based Unscented Transform
IEEE	Institution of Electrical and Electronic Engineering
KCL	Kirchhoff's Current Law
LHS	Latin Hypercube Sampling
MCS	Monte Carlo Simulation
PDF	Probability Density Function
PEM	Point Estimate Method
PLF	Probabilistic Load Flow

PQ bus	Load bus
PV bus	Generator bus
RE	Renewable Energy
T.S.M	Taylor Series based Multivariate conventional Unscented Transform
UDR	Univariate Dimension Reduction
UT	Unscented Transform
VUF	Voltage Unbalance Factor

Chapter 1

Introduction

In this chapter, a brief introduction and the motivation for the research work are discussed. The main objectives of the work and the outline of the thesis are also presented.

1.1. Background and Motivation

Power systems like most other component systems have undergone evolutions which have fast-tracked their development in the last few decades. Recently, most of these developments have been driven by the deregulation of the electricity market alongside the call for sustainable energy practices which has brought renewable energy to limelight globally. These have greatly influenced the high presence of small renewable energy generators within the modern day power system.

Renewable energy generators (such as wind, solar etc.) are dependent on the prevailing weather conditions which are not fixed. Hence the outputs of these generators are unpredictable. This increases the level of uncertainty and risk in the system considering that consumer load demand is also random in nature. The cumulative effects of these uncertainties are of critical importance during power system planning and operations. Load flow studies are often carried out in evaluating the performance of the power system.

In the past, a load flow study was solely deterministic in nature. The need to include the effect of uncertain variables was introduced less than four decades ago. The usual method for uncertainty representation has been based on probability theorem which informed the use of the Monte Carlo Simulation technique as the benchmark [1.1] for some pioneering work on uncertainty based load flow, often referred to as Probabilistic load flow. The Monte Carlo Simulation technique, though accurate, is plagued by very high computational burden such that with many potential variables it can become unwieldy.

Considering this drawback of the conventional Monte Carlo Simulation method, an alternative technique which overcomes this challenge while still possessing the accuracy advantage is desired.

1.2. Objectives of the Thesis

The main objective for this research work is to develop a fast and accurate method for carrying out load flow studies in systems with varying source and or loads. To achieve the stated objective, a novel but simple method which overcomes the computational challenges associated with the Monte Carlo simulation method while having comparable performance in terms of accuracy is desired.

To ensure the developed method is effective in the uncertainty based load flow studies domain, its performance is evaluated against existing methods in terms of speed and accuracy. The developed method should also be realistic and applicable to practical transmission and distribution systems.

1.3. Thesis Outline

The background and motivation for embarking on this research work are discussed in this chapter. The objectives and the thesis outline are also presented.

Chapter 2 discusses the peculiarities of contemporary power systems and the essentiality of power/load flow studies in ensuring proper system performance. Methods for load flow studies are reviewed while the mathematical basis for load flow studies using the Newton Raphson method and the Backward-Forward Sweep technique for transmission and distribution systems respectively are summarised.

In chapter 3, approaches for uncertainty representation are reviewed while also discussing possible sources of uncertainty within the power system. The basics of probability analysis are also presented alongside the probabilistic models for

uncertainty variables within the power system. Results for the validation of wind speed and wind power output using measured data are also shown.

Chapter 4 reviews existing probabilistic load flow techniques while highlighting their strengths and drawbacks. The features of an effective probabilistic load flow method are also outlined.

In chapter 5 the conventional Unscented Transform method [1.2] is introduced as a method for probabilistic load flow studies. The mathematical background for the method is presented and its applicability in probabilistic load flow is demonstrated using a simple 6 bus test system [1.3]. The challenges of using the method in systems with more than one randomly varying parameter and for arbitrary varying distributions (such as the wind power distribution) are outlined.

In Chapter 6, the Unscented Transform is viewed as a Gaussian quadrature problem. Orthogonal polynomials were described for solving for the (sigma) estimation point and weights of random variables. The Stieltjes procedure is introduced as a technique for estimating the sigma points and weights for arbitrary randomly varying distributions. Tensor product is employed in determining final sigma points and weights for system involving more than one random variable. The performance of the technique is evaluated using the 6 bus test system. The problem of "curse of dimensionality" associated with employing the approach for a multivariate problem is identified.

Chapter 7 presents the dimension reduction technique as an approach in resolving the *curse of dimensionality* problem. The performance of the enhanced Unscented Transform method is evaluated for small and large test systems. The performance of the method is also demonstrated for a real practical distribution system.

Chapter 8 details an extension of the enhanced Unscented Transform method in treating dependent variables which are common in real life systems.

Chapter 9 summarises the thesis with the conclusion while also identifying challenges/limitations faced in the research. Potential future areas for further studies are finally highlighted.

1.4. References

[1.1] R.N. Allan, A.M. Leite da Silva and R.C. Burchett "Evaluation Methods and Accuracy in Probabilistic Load Flow Solutions" *IEEE Trans. on Appra and Systems*, PAS-1000, No 5, May 1981

[1.2] S. J. Julier, & J. Uhlmann, "Consistent Debiased Method for Converting Between Polar and Cartesian Coordinate Systems". *In Proc. 1997 SPIE Conference on Acquisition, Tracking, and Pointing*. 3086. SPIE

[1.3] A. J. Wood, and B. F. Wollenberg, *Power Generation, Operation and Control* 2nd Edition, John Wiley & Sons, 1996 pp 123-124

Chapter 2

The Modern Day Power System

The modern day power system has gradually evolved from its usual simple unidirectional flow path to a complex multidirectional flow system. This has been influenced by the high influx of renewable energy systems with unpredictable outputs and deregulated electricity system amongst other factors. In view of this level of uncertainty and added risk, proper monitoring of the system is essential in maintaining continuous operation of the system during both steady state and contingency operating conditions. This chapter discusses power system analysis with emphasis on load/power flow studies since it is the bedrock of most of the other power system analysis. The mathematical formulation and methods for load flow studies in transmission systems and distribution systems are expounded.

2.1. The Peculiarities and Growth of the Modern Day Power System

Power systems like most other systems have experienced growth and huge developmental processes in the last few decades. The initial power system, which was simple and supported unidirectional flow of power from the generator (source) through the transmission (medium) system to the distribution part (receptor) has grown now into a complex multidirectional flow system. This has been credited to the advent of deregulated open access power system and the introduction of decentralised generation amongst other things. Equally important is the global call for sustainable energy practices and the reality of the depletion of fossil fuels, which have helped encourage decentralized generation through renewable energy systems.

Although, opponents of renewable energy systems point to high initial cost, uncertainty in its output and some grid connection issues as the main limit to its acceptability. Given, the global call and commitment to carbon emission reduction, policies/incentives been put in place by various governments have

promoted renewable energy and decentralised generation. Some of these policies/incentives include feed-in-tariffs, electric utility obligation, tradable renewable energy certificate (REC), renewable portfolio standards (RPS), investment/production tax credits, reduced tax and VAT amongst other things [2.1]. Equally, potential advantages such as reduced transmission losses, provision of grid support and improved power quality [2.2, 2.3] if properly managed have promoted the installation of renewable energy generators.

All these have promoted interest and fast growth of the renewable energy sector. For instance, in 2011, 71% of total electric capacity added in the European Union (EU) was from renewable systems, while the US and China had 39% and more than 33% in the same period respectively [2.1]. The EU targets increasing its renewable energy capacity to 20% by 2020 [2.4].

In view of these developments, the distribution part of the power system, which was hitherto seen as the power receptor, now comprises some form of (small and medium scale) renewable generation, which makes power flow sometimes bidirectional. This further adds to the complexities of the modern day system and hence calls for stricter regulations and monitoring of the contemporary power system to ensure supposed inherent advantages are maximised.

To ensure compliance of the power system to stipulated regulations and operating conditions during steady state and or contingency conditions, power systems analysis is often carried out by power engineers to monitor the system performance. The next section discusses power system analysis tools and studies.

2.2. Power System Analysis

An electric power system comprises electrical components that generate, transmit/distribute and consume (load) the electric power. This involves various types of electrical equipment such as generators, transformers, reactors, capacitors, transmission lines, circuit breakers etc. Power system analysis is concerned with the understanding of the operation of these various complex components as a whole unit [2.5]. To fully understudy power systems, analysis

is carried out both during steady state and transient or dynamic operating condition of the system [2.6]. In steady state analysis, all transients/disturbances are assumed to have settled down while analysing the performance of the system. The power flow (load flow) study is the major tool used for steady state analysis [2.5-2.7]. Power stability analysis, short circuit analysis, motor start analysis and fault analysis are typical examples of studies carried out for transient and dynamic system states [2.6, 2.7]. Harmonic studies are also carried out to understand the impact of harmonics on the system and evaluate the performance of designed filters for curtailing identified effects.

Although, all the above stated analyses are important to power systems engineers and essential in ensuring the overall wellbeing of the power network, power (load) flow analysis is the most popular of the analyses [2.7, 2.8] and often provides the starting point for other analysis. The importance and basic principles behind load/power flow studies is discussed in the following section.

2.3. Load (Power) Flow Analysis

Load (power) flow study is carried out to determine the performance of the system during steady state and contingency operating conditions. During power flow studies, the non-linear power flow equations relating the various components are solved from the generator through the transmission/distribution link to the load. This is done with the aim of evaluating the voltage magnitude and angle on each busbar and the active and reactive power flows through the line. The information amongst other things helps to [2.5-2.9];

i. Ensure the limits (line and voltage) at the consumer side and the system as a whole are strictly adhered to
ii. Determine the power losses within the system
iii. Determine the proper setting for transformers within the system
iv. Identify the need for extra generators and the right location for their installation.
v. Locate capacitor and reactors to reduce losses within the system
vi. Study the effect of generator loss or other contingencies

vii. Identify and evaluate the effect of possible or proposed improvement to the system

In all, it is seen that load flow analysis is an essential decision-making tool during planning, operation, control and expansion stages of a power system to ensure proper balance of the network.

The power system is made up of three distinct types of buses; generator bus, load bus and slack bus. Generator buses (PV bus) are connection points for generators within the system that have known voltage magnitude and active power while the active and reactive power for the load buses (PQ bus) are usually specified. The slack bus (swing bus) is the fictitious reference point for the system, which is usually a generator able to supply or consume power. The voltage magnitude and angle of the slack bus must be known (reference) before the load flow problem can be solved.

A load flow study is often carried out both in the transmission part of the network and the distribution part of the network. The difference between these two and how the study is performed will be described.

In any load flow study, accuracy, speed, reliability, simplicity, versatility, and low computer storage space are important factors [2.8]. All load flow methods or techniques are therefore designed and rated using these properties as the adjudicating factor in determining their adequacy.

2.3.1. Load (Power) Flow Analysis in Transmission Systems

The transmission system in the context of transmission system load flow analysis comprises the whole power system aside from the distribution system. The load connected to each major busbar or substation is viewed as a single entity without considering its further distribution through the radial or meshed structure. These systems are therefore simple, straightforward and balanced thus allowing for easy computation of power flow in them. The system is represented using a single-phase line since it is balanced. In view of this, most of the initial load flow solutions have been based on the phase-balance

assumption, which leads to convergence problems and significant errors in ill-conditioned systems like the distribution systems (next section).

2.3.1.1. Load Flow Formulation

The transmission system load flow formulation can be expressed as follows. Let the voltage of an n bus power system be represented by matrix vector **V** with bus admittance matrix **Y**.

$$\mathbf{V} = \begin{bmatrix} |V_1|\angle\partial_1 = |V_1|(\cos\partial_1 + j\sin\partial_1) \\ \vdots \\ |V_n|\angle\partial_n = |V_n|(\cos\partial_n + j\sin\partial_n) \end{bmatrix} \tag{2.1}$$

$$\mathbf{Y} = [\mathbf{G} + j\mathbf{B}] \tag{2.2}$$

Following Kirchhoff's Current Law (KCL), the current injected into a bus *i* is (2.3);

$$\mathbf{I_i} = \mathbf{Y_{i1}V}_1 + \mathbf{Y_{i2}V}_2 + \cdots + \mathbf{Y_{in}V_n} \tag{2.3}$$

This can be simplified as;

$$\mathbf{I_i} = \sum_{k=1}^{n} \mathbf{Y_{ik}V_k} \tag{2.4}$$

The active and reactive power flowing to bus *i* is given by (2.5).

$$P_i + jQ_i = V_i I_i^* \tag{2.5}$$

where I_i^* is the complex conjugate of the current.

Substituting (2.4) into (2.5) results into (2.6).

$$P_i + jQ_i = V_i \sum_{k=1}^{n} (Y_{ik}V_k)^* \tag{2.6}$$

Y_{ik} can also be written as;

$$Y_{ik} = |Y_{ik}|\angle\theta_{\mathrm{ik}} = |Y_{ik}|(\cos\theta + j\sin\theta) = \mathrm{G} + j\mathrm{B} \tag{2.7}$$

Rewriting (2.6) in polar form gives (2.8).

$$P_i + jQ_i = V_i \sum_{k=1}^{n} |V_k Y_{ik}| \exp\ j(\partial_i - \partial_k - \theta_{ik}) \quad (2.8)$$

Equating the real and imaginary parts of the above equation gives two equations, which are the equivalent active and reactive power respectively.

$$P_i = V_i \sum_{k=1}^{n} |V_k Y_{ik}| \cos(\partial_i - \partial_k - \theta_{ik}) \quad (2.9)$$

$$Q_i = V_i \sum_{k=1}^{n} |V_k Y_{ik}| \sin(\partial_i - \partial_k - \theta_{ik}) \quad (2.10)$$

Considering that the sum of the power at a point should be zero, (2.9) and (2.10) can be written as;

$$P_{Gi} - P_{Li} = P_i = V_i \sum_{k=1}^{n} |V_k Y_{ik}| \cos(\partial_i - \partial_k - \theta_{ik}) \quad (2.11)$$

$$Q_{Gi} - Q_{Li} = Q_i = V_i \sum_{k=1}^{n} |V_k Y_{ik}| \sin(\partial_i - \partial_k - \theta_{ik}) \quad (2.12)$$

Applying the above equation for all the buses excluding the slack bus results in 2(n-1) simultaneous equations with P_i, Q_i, $|V_i|$ and δ_i, as unknowns depending on the bus type. These equations are referred to as the power flow equations.

2.3.1.2. Load Flow Methods

Several methods mostly based on iteration have been proposed in the past for solving the load flow equations presented above. One of the earliest methods was based on the Gauss-Seidel method while the Newton Raphson method was later introduced due to the slow convergence of the former amongst other challenges. A brief comparison of the two techniques is presented in Table 2.1 below [2.5-2.9].

Table 2.1: Comparison of the Newton Raphson and Gauss-Seidel Methods

Feature	Newton Raphson	Gauss-Seidel
Convergence rate	Quadratic convergence. Fewer iterations for convergence	Linear convergence rate
Number of Iteration	Independent of bus (size) number	Increases linearly as number of bus (size) increases

Size	Suitable for large sized systems	Suitable for small sized systems
Storage space	More memory space	Less memory space
Simplicity	Jacobian matrix computed during each iteration thus requires more computation per iteration.	Requires lower number of computation per iteration
Slack bus	Convergence independent of slack bus selection	Convergence affected by slack bus selection

Most power systems are relatively large, this amongst other things like better convergence makes the Newton Raphson method well favoured above the Gauss-Seidel method for load flow studies. As such, the Newton Raphson method is commonly used and it is further discussed.

2.3.1.3. Newton Raphson Method

The Newton Raphson method is a well-established iterative method which is used for solving for the unknown in a set of simultaneous equations. In load flow analysis, the set of non-linear equations (2.11, 2.12) is approximated using a set of initially guessed values, as a set of linear equations by employing the Taylor series expansion with first order truncation [2.9].

The mathematical formulation of the Newton-Raphson method with reference to load flow studies is discussed below starting with (2.11) and (2.12).

With the Newton Raphson method, the main aim is to reduce the mismatch between the real and calculated value to zero. That is, the mismatch between the calculated P_i and Q_i in (2.11) and (2.12) can be written as;

$$P_{Gi} - P_{Li} - P_{i(calculated)} = \Delta P_i \tag{2.13}$$

$$Q_{Gi} - Q_{Li} - Q_{i(calculated)} = \Delta Q_i \tag{2.14}$$

The relationship between the change (mismatch) in power (active and reactive) and change in voltage magnitude and angle is given by (2.15).

$$\begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix} \begin{bmatrix} \Delta \delta \\ \Delta V \end{bmatrix} \tag{2.15}$$

For a *n*-bus system with n_g generator buses (PV buses) and n_l load buses (PQ buses) with bus 1 assumed to be the slack bus, the formation of the Jacobian matrix sub-matrices is detailed below.

1. Formation of J_{11}

The sub-matrix J_{11} reflects the change in the active power relative to the voltage angle change. The slack bus voltage angle is fixed as such, its effect is ignored.

$$J_{11} = \begin{bmatrix} \frac{\partial P_2}{\partial \delta_2} & \cdots & \frac{\partial P_2}{\partial \delta_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial P_n}{\partial \delta_2} & \cdots & \frac{\partial P_n}{\partial \delta_n} \end{bmatrix} \tag{2.16}$$

The matrix J_{11} is a (*n-1)* x (*n-1*) matrix whose elements are derived by differentiating (2.9) with respect to δ, the voltage angle. The diagonal and off-diagonal elements are given as (2.17) and (2.18) respectively.

$$\frac{\partial P_i}{\partial \delta_i} = -V_i \sum_{k=1,\, k \neq i}^{n} |V_k Y_{ik}| \sin(\delta_i - \delta_k - \theta_{ik}) \tag{2.17}$$

$$\frac{\partial P_i}{\partial \delta_k} = |V_i V_k Y_{ik}| \sin(\delta_i - \delta_k - \theta_{ik}) \;\; i \neq k \tag{2.18}$$

2. Formation of J_{12}

This sub-matrix is formed by differentiating the active power with respect to the voltage magnitude (2.19). Since the voltage magnitudes of the generator buses are fixed, their derivatives are neglected. The matrix thus contains (*n-1*) x (n_l) elements. The elements are derived using (2.20) and (2.21) for the diagonal and off-diagonal terms respectively.

$$J_{12} = \begin{bmatrix} \frac{\partial P_2}{\partial |V_2|} & \cdots & \frac{\partial P_2}{\partial |V_n|} \\ \vdots & \ddots & \vdots \\ \frac{\partial P_n}{\partial |V_2|} & \cdots & \frac{\partial P_n}{\partial |V_n|} \end{bmatrix} \tag{2.19}$$

$$\frac{\partial P_i}{\partial |V_i|} = 2V_i\ Y_{ii}\cos(\theta_{ii}) + \sum_{k=1,k\neq i}^{n} |V_k Y_{ik}| \cos(\delta_{\mathrm{i}} - \delta_{\mathrm{k}} - \theta_{\mathrm{ik}}) \tag{2.20}$$

$$\frac{\partial P_i}{\partial |V_k|} = |V_i Y_{ik}| \cos(\delta_{\mathrm{i}} - \delta_{\mathrm{k}} - \theta_{\mathrm{ik}})\ \ i \neq k \tag{2.21}$$

3. Formation of J_{21}

The sub-matrix J_{21} is formed by components which are partial derivatives of the reactive power with respect to the voltage angle (2.22). Using (2.22), the diagonal and off-diagonal elements of the sub-matrix are expressed in (2.23) and (2.24) respectively. In all, the matrix is a (n_l) x (n-1) matrix.

$$J_{21} = \begin{bmatrix} \frac{\partial Q_2}{\partial \delta_2} & \cdots & \frac{\partial Q_2}{\partial \delta_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial Q_n}{\partial \delta_2} & \cdots & \frac{\partial Q_n}{\partial \delta_n} \end{bmatrix} \tag{2.22}$$

$$\frac{\partial Q_i}{\partial \delta_i} = V_i \sum_{k=1,k\neq i}^{n} |V_k Y_{ik}| \cos(\delta_{\mathrm{i}} - \delta_{\mathrm{k}} - \theta_{\mathrm{ik}}) \tag{2.23}$$

$$\frac{\partial Q_i}{\partial \delta_k} = -|V_i V_k Y_{ik}| \cos(\delta_{\mathrm{i}} - \delta_{\mathrm{k}} - \theta_{\mathrm{ik}})\ \ i \neq k \tag{2.24}$$

4. Formation of J_{22}

The elements of the last matrix block are formed based on the derivative of the reactive power with respect to the voltage magnitude. The matrix J_{22} is given by (2.25).

$$J_{22} = \begin{bmatrix} \frac{\partial Q_2}{\partial |V_2|} & \cdots & \frac{\partial Q_2}{\partial |V_n|} \\ \vdots & \ddots & \vdots \\ \frac{\partial Q_n}{\partial |V_2|} & \cdots & \frac{\partial Q_n}{\partial |V_n|} \end{bmatrix} \tag{2.25}$$

The diagonal and off-diagonal elements of the matrix are;

$$\frac{\partial Q_i}{\partial |V_i|} = 2V_i Y_{ii} \sin(\theta_{ii}) + \sum_{k=1,k\neq i}^{n} |V_k Y_{ik}| \sin(\delta_i - \delta_k - \theta_{ik}) \tag{2.26}$$

$$\frac{\partial Q_i}{\partial |V_k|} = |V_i Y_{ik}| \sin(\delta_i - \delta_k - \theta_{ik}) \quad i \neq k \tag{2.27}$$

The matrix J_{22} is made up of (n_l) x (n_l) elements. In all, the Jacobian matrix is a (n+n_l-1) x (n+n_l-1) matrix. A summary of the Newton Raphson algorithm for the load flow problem is presented below using the flowchart in Fig. 2.1.

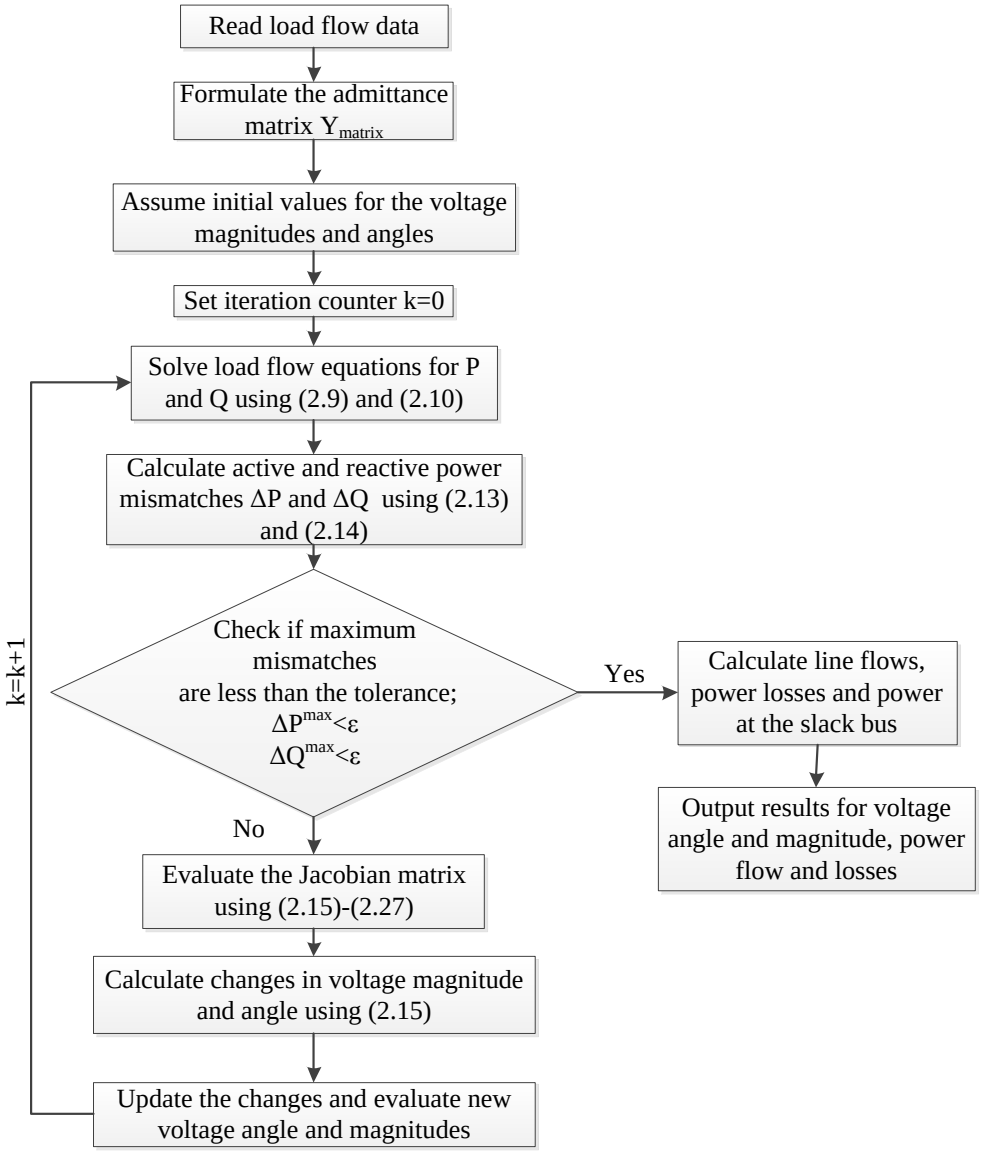

Fig 2.1: Load Flow Studies Using the Newton-Raphson Method

Also worthy of mention is the Fast Decoupled Load Flow (FDLF) method, which approximates the Newton Raphson method. The approximation exploits the loose coupling between active power (and voltage angle) and the voltage magnitude (and reactive power). As such, a change in the active power (or voltage angle) is assumed to have no significant effect on the voltage magnitude (or reactive power) and vice versa, since they are loosely coupled, as such it can be neglected. This reduces the computational requirement since sub-matrices J_{12} and J_{21} are zero. The modified Jacobian matrix becomes;

$$\begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} = \begin{bmatrix} J_{11} & 0 \\ 0 & J_{22} \end{bmatrix} \begin{bmatrix} \Delta \delta \\ \Delta \mathrm{V} \end{bmatrix} \tag{2.28}$$

The Jacobian matrix is only calculated at the start of the computation unlike being calculated during each iteration for the generic Newton Raphson method. The FDLF method thus performs better than the generic Newton Raphson in terms of speed and storage space requirements [2.5, 2.9]. However, since the Jacobian matrix is calculated once, this introduces some errors, which requires more iteration to reach the same level of accuracy as the generic method. The quadratic convergence property is also lost at this point. For lines with high resistance to reactance ratio, errors due to this assumption become overly large thus making the technique unreliable [2.10]. In summary, it is seen that the FDLF method is useful in making approximate fast estimations, the full Newton Raphson method is better suited [2.7, 2.10].

2.3.2. Load Flow Studies in Distribution Systems

The modern day distribution system in contrast to a traditional one comprises both generation and load in the system. This makes the concept of load flow in distribution systems even more challenging since they now have to cope with bidirectional power flow and a higher level of uncertainties. Distribution systems have unique features which distinguish them from transmission systems. Some of these features include [2.11-2.14];

i. Radial or a weakly meshed nature

ii. High resistance to reactive impedance ratio (R/X)

iii. Unbalanced operation state

iv. Multiphase

v. Dispersed generation

vi. Large number of branches and nodes

Considering some of the assumptions, the traditional Newton-Raphson method and the Fast decoupled method see the distribution system as "ill-conditioned" hence they result in poor convergence [2.11]. In fact, it was observed to diverge in [2.11] in most of the cases considered. In the same vein, the conventional Gauss-Seidel method is extremely inefficient in solving large power systems, which is a typical feature of distribution systems [2.11]. In general, load flow techniques employed in transmission systems cannot be directly applied to distribution systems.

In view of the shortcomings of the classical methods (for transmission systems), other methods have been proposed to efficiently solve the distribution load problem while fully considering its peculiarities. Several methods [2.11-2.30] have been proposed for carrying out load flow studies in distribution systems. Generally, most of these methods are based on backward-forward substitution, modified Newton Raphson, and Gauss approach. However, most of the methods require some form of backward-forward substitution in attaining convergence.

The Gauss Z_{bus} approach is one of the earliest techniques used for the distribution load flow problem [2.15]. The approach uses the Y_{bus} matrix (which is sparse) and current injections to solve the network equations. The Z_{BR} approach was introduced in [2.16]. Unlike the previous Y_{bus} technique, it utilizes the branch impedance matrix, the current injection alongside a branch-path incidence matrix in solving the network equations [2.16].

Improved Newton Raphson methods have also been employed for load flow studies [2.17-2.20]. In [2.17] and [2.18], a Newton Raphson method with a reduction in the Jacobian matrix (to save time) was proposed. A current injection based on Newton Raphson method is employed in [2.19] for the load

flow studies. In [2.21] a sequence based load flow using Newton Raphson was proposed for solving load flow in unbalanced distribution systems. One drawback of the Newton Raphson based technique is the need to evaluate the Jacobian matrix during each iteration process [2.9]. In addition, its inherent convergence problem still occurs for systems with high unbalance and loading [2.11].

With the Backward-Forward sweep (BFS), the current or power injected into a branch is computed relative to the end voltage while the voltage drop is evaluated in the forward sweep using the calculated current and or power. In [2.11, 2.12], the Backward-Forward sweep method with a layered structured branch numbering is employed. However, because the method can only cater for radial distributions, it was extended to the weakly meshed network in [2.11] using the breaking point mechanism. The ladder network, as employed in [2.13] is based on BFS. The models for representing various transformers, voltage regulators, capacitors, etc. are detailed in [2.13]. A technique called the fast and flexible radial power flow (FFRPF), which is a variant of BFS, is used in [2.14]. Here, bus and branch related matrices are formed and used in updating the voltages and currents. BFS methods with improvements (e.g. in data handling, bus numbering etc.) are also proposed in [2.22-2.24]. All these techniques are derivative free unlike the Newton Raphson based methods; rather, they employ simple circuit analysis which makes them attractive.

Other methods such as the direct approach [2.25-2.26] have also been previously introduced. With the direct approach, two main matrices are required; the first relates the branch current to the load current while the other matrix reflects the relationship between the branch current and the bus voltage. Although the technique is simple, fast and straightforward, it does not accommodate unbalanced or distributed load [2.25].

Of the techniques mentioned, back-forward schemes are mostly used due to their simplicity and non-requirement of a Jacobian matrix, amongst other factors. This informs its use as the benchmark for other distribution system load flow methods [2.16]. An algorithm based on the BFS scheme is discussed further in the succeeding section.

2.3.2.1. The Backward Forward Sweep

The procedure for the distribution load flow described below is adapted from the back-forward sweep and the ladder network as presented in [2.13] and [2.27]. The current is evaluated in the backward sweep while the voltage is updated in the forward sweep. The method is based on child-parent branch identification fully described in this section. The procedures involved are explained in the steps below.

Step 1: Formulate a parent child matrix to identify the connected nodes before and after a node using the network information. For instance, considering a simple 8 bus test system [2.18] shown in Fig. 2.2, the parent child matrix *Mpc* is given by Table 2.2. The *Mpc* matrix is very easy to form as it does not involve node renumbering, the only clear regulation is that the branch number should be one less than the value of the receiving node number. The *Mpc* matrix is very important for the forward and backward sweep since it gives the sending and receiving nodes.

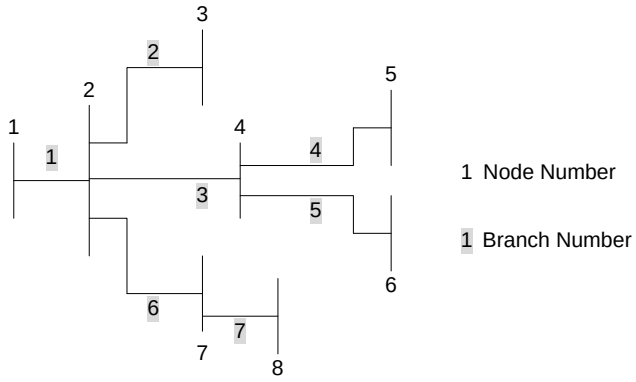

Fig 2.2: A Simple 8 Bus Distribution Test System

Table 2.2: The M_{PC} Matrix for the 8 Bus Test System

Node	Parent	Children
1	-	2
2	1	3,4,7
3	2	-
4	2	5,6

5	4	-
6	4	-
7	2	8
8	7	-

Step 2: Calculate the total nodal current from the loads, shunt capacitor, distributed generation etc. These currents are calculated depending on the type of load (constant impedance, constant current and constant power) and the load connection type (Delta or Wye). Detailed explanations are found in [2.13] and [2.27].

Step 3: The backward sweep is carried out to update the current in the branches working from the end node towards the source node. A single line diagram of the distribution line segment is shown in Fig. 2.3. The update formula used is dependent on the connector (distribution line, transformer or switch) between the nodes. The update formulae for nodes connected with distribution lines are given in (2.29)-(2.32), while (2.33) and (2.34) are respectively for nodes connected by switches or transformers [2.13]. The values for the coefficients c_t and d_t in (2.3.4) vary depending on the transformer configuration. These values are given in Appendix A for wye-wye and delta-wye transformers, values for other configuration can be found in [2.13].

Distribution Lines:

$$I_{kp} = I_{GCLk} + \sum I(children\ \text{branches})_{\text{k}} \tag{2.29}$$

$$I_k = c[V_k] + d[I_{kp}] \tag{2.30}$$

$$c = [Y_k] + \frac{1}{4}[Z_k][Y_k]^2 \tag{2.31}$$

$$d = [U] + \frac{1}{2}[Z_k][Y_k] \tag{2.32}$$

Switches and Transformers:

$$I_k = I_{kp} \tag{2.33}$$

$$I_k = c_t[V_k] + d_t[I_{k+1}] \tag{2.34}$$

where I_{GCL} denotes the sum of current injected by the (embedded/distributed) generators, capacitors and loads on node *k*. I_{kp} gives the nodal current on bus *k*

plus currents from children branches, I_k represents the branch current for line k, V_k is the voltage on node k, Z_k and Y_k are respectively the impedance and admittance matrix for line k, c_t and d_t are coefficients dependent on the transformer configuration and U is a unit matrix.

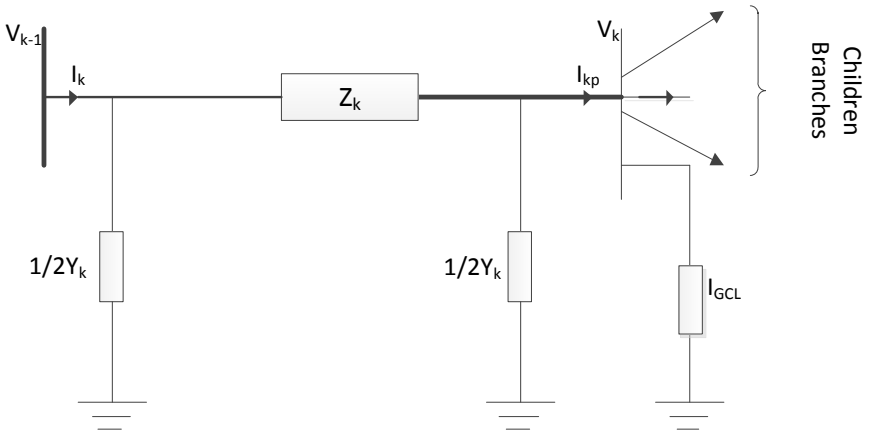

Fig 2.3: Single Line Diagram of the Distribution Line Segment

Step 4: In the forward sweep, the main aim is to evaluate the voltage drop across the lines in order to estimate the voltage at each of the nodes. The voltage is evaluated starting from the source node and gradually moving towards the end node. This is evaluated using (2.35) [2.13].

$$V_k = [A][V_{k-1}] - [B][I_k] \tag{2.35}$$

The parameters A and B are determined based on the type of connection between the nodes. The values of A and B for grounded wye-grounded wye and delta-grounded wye transformers are presented in Appendix A while those for connection types can be found in [2.13]. For nodes connected by switches, a zero drop is assumed across the switch, hence $V_k = V_{k-1}$.

Step 5: Once the backward and forward sweeps are completed, the convergence test is carried out to ensure the errors are within the limits. The iteration is terminated once the errors between the voltage (on all phases) from the previous iteration and the current iteration are within the prescribed limits.

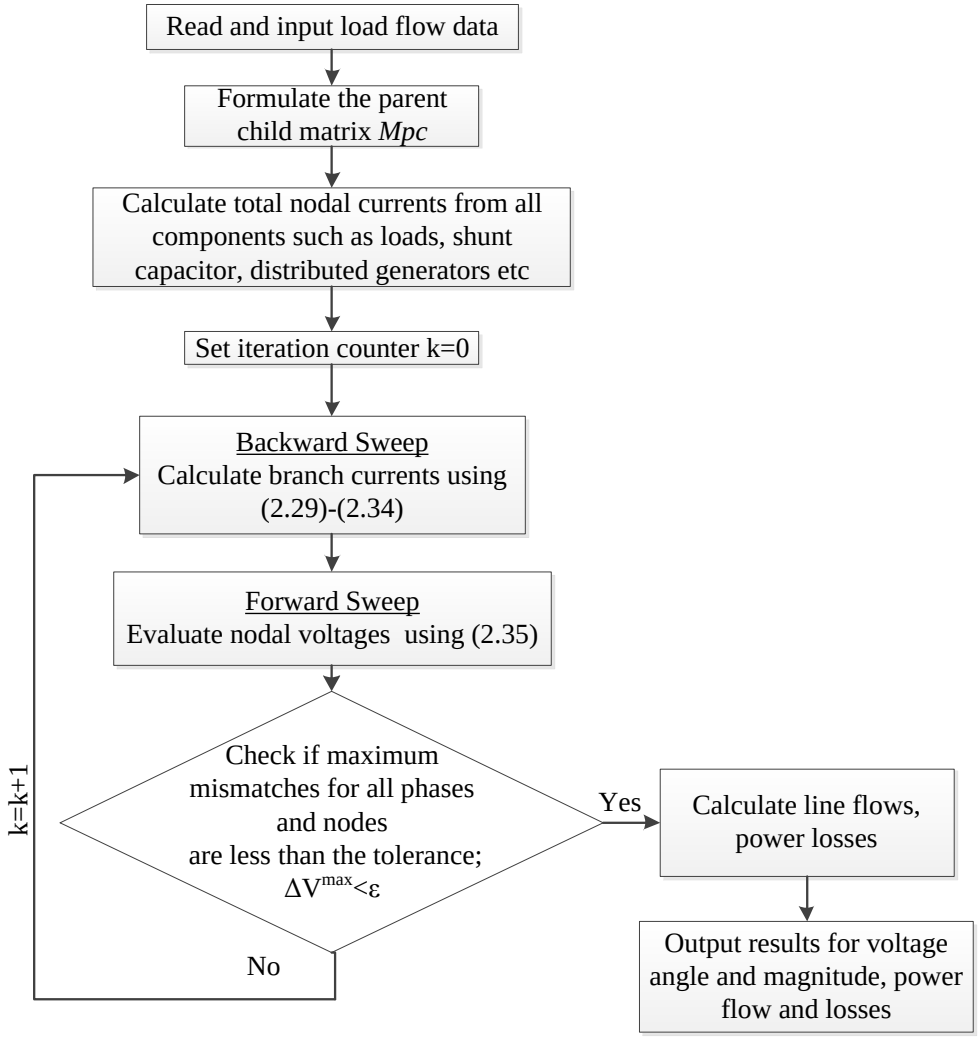

Fig 2.4: Backward-Forward Sweep for Distribution Load Flow Study

A summary of the load flow process using the backward-forward sweep method is presented in Fig. 2.4.

2.3. Shortfalls of Deterministic Load flow Studies

The methods for load flow studies for both transmission and distribution systems have been discussed. The methods are seen to be accurate in solving the load flow problem, however, the mathematical formulations neglect the effects of uncertainty, which are prevalent in power systems. This is even more important considering the high influx of non-dispatchable renewable energy systems whose outputs are most dependent on weather conditions.

This shortfall was initially identified and treated by Borkowska in her research publication of 1974 [2.28]. In the work, the effects of uncertainties due to the varying nature of load demands and generator outputs were viewed to be probabilistic in nature, thus probabilistic mathematical models were employed. Since this initial proposition, several research work have focused on appropriate techniques/approaches to represent and or model the effects of uncertainties while carrying out load flow studies. This is important in ensuring that the true system state is known and well represented, considering that the outputs of load flow analysis influences decisions made by power engineers.

2.4. Précis

The modern day power system is advanced in its makeup and operations. This is largely due to the high penetration level of renewable energy systems, deregulated electricity system and the global call for sustainable energy practices. This is seen to present some benefits such as improved power quality, efficiency and grid support amongst other things. To fully harness these benefits, proper management of the power system is essential. This can be done using the numerous power system analysis studies discussed in this chapter. Although several analysis techniques such as short-circuit analysis, harmonic studies, power system stability, power/load flow studies are normally carried out to ensure overall healthy operation of the power system, power/load flow studies is the most popular since other studies build on it.

The importance of power/load flow studies have been discussed in this chapter and the mathematical formulation. The features, advantages and disadvantages of the two main load flow studies method namely the Newton Raphson and Gauss-Seidel method have also been examined. Of the two, the Newton Raphson is more attractive due to its good convergence properties and applicability to large systems which are very common. The detailed process for implementing the Newton Raphson method for load flow studies is further presented. Also mentioned is the Fast Decoupled Load flow method, which builds on the Newton Raphson technique.

The peculiarities of the distribution system and the convergence problems (amongst other things) associated with using transmission load flow techniques for distribution load flow problems have been explained. Techniques such as improved Newton Raphson, direct method, ladder method, Gauss approach (Z_{bus} and Z_{BR}) and Backward-Forward sweep were briefly discussed with the aim of understanding their strengths and weaknesses. Of the methods, a scheme based on Backward-Forward sweep was further discussed due to its good adaptability for the distribution load flow problem, its mathematical formulation is also presented.

Considering the high level of uncertainties and risks introduced to the power system by renewable systems and the variability in load demand, their effects cannot be ignored during load/power flow studies if the true system state is desired. The methods earlier used for load flow studies both in transmission and distribution systems have been deterministic in nature. The concept of accounting for uncertainties during load flow studies was mentioned with reference to the initial proposition on probabilistic load flow. Since this initial proposition, several other research works aimed at understanding uncertainty representation and developing good methods for uncertainty based load flow studies have been carried out. This forms the basis for the succeeding chapters. The various approaches for representing uncertainties are discussed in the next chapter.

2.5. References

[2.1] RE21, Renewable Energy 2012, Global Status Report [Online]. Available: www.map.ren21.net/GSR/GSR2012_low.pd [accessed: 19/12/2012].

[2.2] G. Pepermans, J. Driesen, D. Haeseldonckx, W. D'haeseleer and R. Belmans, "Distributed Generation: Definition, Benefits and Issues" Working Paper No 2003-8, Energy Transport and Environment, Katholieke Universiteit Leuven, August 2003 [Online]. Available: http://www.econ.kuleuven.ac.be/ew/academic/energmil/downloads/ete-wp-2003-08.pdf [accessed: 19/12/2012].

[2.3] Distributed Generation Primer, Report prepared for the U.S. Department of Energy, (DOE/NETL-2002/1174), [Online]. Available: http://www.casfcc.org/2/StationaryFuelCells/PDF/Distributed%20Generation%20Primer.pdf [accessed: 19/12/2012].

[2.4] The EU's Target for Renewable Energy: 20% by 2020, 27th Report of Session 2007-2008, European Union Committee, (Volume I, HL Paper No 175-I) [Online]. Available: http://www.publications.parliament.uk/pa/ld200708/ldselect/ldeucom/175/175.pdf [accessed: 19/12/2012].

[2.5] P. S. R. Murthy, *Power System Analysis,* BS Publications, Hyderabad, India, 2007 pp. 98-175

[2.6] "Power Systems Analysis," [Online]. Available: http://www.kalkitech.com/solutions/power-system-analysis [accessed: 13/12/2012]

[2.7] M. Venkatasubramanian and K. Tomsovic, "Power Systems Analysis," *Electrical Engineering Handbook*, Elsevier Academic Press, 2005, pp. 779-785

[2.8] B. Stott "Review of load flow calculation methods," *In proc. of IEEE,* vol. 62, no.7, pp. 916-929 July 1974

[2.9] N. Kumar and S. Kumar, *Power System Analysis*, Asian Books Private Limited, New Delhi, India, 2010 pp.150-202

[2.10] I. A. Hiskens, Notes on "Power Flow Analysis," University of Wisconsin-Madison, November 2003

[2.11] D. Shirmohammadi, H. W. Hong, A. Semlyen, and G. X. Luo, "A compensation Based Power Flow Method for Weakly Meshed Distribution and Transmission Networks," *IEEE Trans. Power Systems*, vol. 3, no. 2, pp. 671-679, May 1988

[2.12] C. S. Cheng, and D. Shirmohammadi, "A three phase Power flow method for real time distribution system analysis," *IEEE Trans. Power Systems*, vol. 10, no. 2, pp. 671-679, May 1995

[2.13] W. H. Kersting, *Distribution System Modelling*, CRC Press, LLC 2006

[2.14] M. F. AlHajri, and M. E. El-Hawary, "Exploiting the Radial Distribution Structure in Developing a Fast and Flexible Radial Power

Flow for Unbalanced Three Phase Networks," *IEEE Trans. on Power Delivery*, vol. 25, no. 1, pp. 378-389, Jan. 2010

[2.15] T. H. Chen, M. S. Chen, K. J. Hwang, P. Kotas, and E. A. Chebli, "Distribution System power flow analysis- A rigid approach," *IEEE Trans. Power Delivery*., vol 6, no. 3, pp. 1146-1152, July 1991

[2.16] T. H. Chen, and N. C. Yang, "Three phase power flow by direct Z_{BR} method for unbalanced radial distribution systems," *IET Generation Transmission & Distribution*, vol. 3, iss. 10, pp. 903-910, 2009

[2.17] R. D. Zimmerman, and H. D. Chiang, "Fast Decoupled Power Flow for Unbalanced Radial Distribution Systems," *IEEE Trans. Power Systems*, vol. 10, no. 4, pp. 2045-2052, Nov. 1995

[2.18] J. H. Teng, and C. Y. Chang, "A Novel and Fast Three Phase Load Flow for Unbalanced Radial Distribution Systems," *IEEE Trans. Power Systems*, vol. 17, no. 4, pp. 1238-1244, Nov. 2002

[2.19] P. A. N. Garcia, J. L. R. Pereira, S. Carneiro, V. M. da Costa and N. Martins, "Three-Phase Power Flow calculations using the current injection method," *IEEE Trans. Power Systems*, vol. 15, no. 2, pp. 508-514, May 2000

[2.20] W-M. Lin., Y-S. Su, H-C. Chin and J-H. Teng, "Three-Phase Unbalanced Distribution Power Flow Solutions with Minimum Data Preparation," *IEEE Trans. Power Systems*, vol.14, no.3, pp. 1178-1183, August 1999

[2.21] M. Abdel-Akher, K. M. Nor, and A. H. A. Rashid, "Improved Three Phase Power Flow Methods Using Sequence Components," *IEEE Trans. Power Systems*, vol. 20, no. 3, pp. 1389-1397, Aug 2005

[2.22] E. Janecek, and D. Georgiev, "Probabilistic Extension of the Backward/Forward Load Flow Analysis Method," *IEEE Trans. Power Systems*, vol. 27, pp. 695-704, May 2012

[2.23] F. J. Ruiz-Rodriguez, J. C. Hernandez, and F. Jurado, "Probabilistic load flow for photovoltaic distributed generation using the Cornish Fisher expansion," *Electric Power Systems Research*, Vol. 89, 2012, pp. 129-138

[2.24] S. Mishra, "A Simple Algorithm for Unbalanced Radial Distribution System Load Flow," TECON 2008, IEEE conference for Region 10, 19-20 Nov. 2008

[2.25] J. H. Teng, "A Direct approach for distribution system load flow solutions," *IEEE Trans. Power Delivery*, vol. 18, no. 3, pp. 882-887, July 2003

[2.26] H. E. Farag, E. F. El-Saadany, R. E. Shatshat, and A. Zidan, "A Generalized Power Flow Analysis for Distribution Systems with High Penetration of Distribution Generation," *Electric Power Systems Research*, vol. 81, pp. 1499-1506, 2011

[2.27] R. D. Zimmerman, "Comprehensive distribution power flow: Modelling, formulation, solution algorithms and Analysis," PhD. Thesis, Cornell University, 1995

[2.28] B. Borkowska, "Probabilistic load flow," *IEEE Trans. Power Apparatus and System*, vol. PAS-93, no. 3, pp 752-755, May-Jun, 1974

Chapter 3

Uncertainty Characterisation and Modelling

To properly account for the effect of uncertainties within any given model, the mathematical basis of uncertainties needs to be understood. In this chapter, some mathematical representations of uncertainties are discussed. The probabilistic model of some uncertainties within the modern day power systems are also presented while the model for the output wind power is validated using real measured data.

3.1. Uncertainty Classification

Power systems like most practical systems are not immune to uncertainties. Uncertainties within the power system can be due to natural variability, data uncertainty, measurement errors, human errors and model or method error amongst other factors. Uncertainties have been previously classified using different nomenclatures; the division based on aleatory and epistemic uncertainties is discussed below.

3.1.1. Aleatory Uncertainties

Aleatory uncertainties are due to unpredictable and unbiased variations or natural randomness in a quantity or system. They are also referred to as natural variability, random uncertainty or irreducible uncertainty [3.1-3.6]. In general, they reduce the precision of the output result and cannot be reduced by better measurement or control on the part of the experimenter. With good understanding of the theories associated with them, their effect can be statistically quantified and taken into account while formulating the system /model. One typical example of this is wind speed and solar insolation. This type of uncertainty is the basis for the formulation of probabilistic load flow studies.

3.1.2. Epistemic Uncertainties

Epistemic uncertainties are uncertainties resulting from the process of modelling a system. It is usually caused by lack of knowledge, inexactness, indeterminacy, ignorance, immeasurability, lack of observation and conflicting evidences amongst other things [3.1-3.5]. In load flow studies for instance, this can result from using an overly simplified form of the non-linear load flow equation, model/method inadequacy or ignoring some terms to reduce complexity of the problem. These types of uncertainties affect the level of accuracy of the output. The level of uncertainty can be reduced with a better level of knowledge or information. In fact, the main aim of this work is to reduce the effect of epistemic uncertainty through the proposition of an accurate method for probabilistic load flow studies. The epistemic uncertainty is also referred to as systematic uncertainty [3.4].

It is worth stating that uncertainty due to a random load can sometimes be referred to as aleatory or irreducible uncertainty from the generator's point of view, however with mechanisms such as load scheduling, load control or demand response the effect of load uncertainty can be controlled.

3.2. Representation of Uncertain Variables

Uncertainties have been previously represented using various theories. Some of these theories include the classical set theory, rough set theory, theory of evidence, interval mathematics, fuzzy set theory (possibility theory) and probability theory [3.2]. The last three (interval, fuzzy set and probability) of these theories have been previously applied in representing uncertainties in load flow studies and will be discussed in details below.

3.2.1. Interval Mathematics

Interval mathematics is used in representing uncertainties due to non-specificity or imprecision [3.2]. One major advantage of the interval mathematics is that the information about the distribution of the uncertain

variable is not required since it is based on boundaries. As such, for problems where only the range of existence is known, the method gives a good representation of the uncertainty.

As an illustration, for two intervals A= [a, b] and B= [c, d], the simple interval mathematics operation is [3.8];

$$
\begin{aligned}
&[a,b]+[c,d]=[a+c,b+d]\\
&[a,b]-[c,d]=[a-c,b-d]\\
&[a,b].[c,d]=[\min(ac,ad,bc,bd),\max(ac,ad,bc,bd)]\\
&[a,b]/[c,d]=[a,b].[1/d,1/c] \qquad \text{if} \qquad 0\notin[c,d]
\end{aligned}
\tag{3.1}
$$

One of the drawbacks of the theory is that it does not obey the distributive law [3.7, 3.10] and can result in some complex evaluation for non-real number intervals as with the case of power. In addition, full information about the output distribution cannot be obtained with interval mathematics since only the boundaries are given [3.2]. Another disadvantage of theory is its overestimation of output when dependency exists among variables. This is often referred to as "interval dependency problem" [3.9].

3.2.2. Fuzzy Set Theory

The concept of fuzzy set was developed by Zadeh for uncertainties which are not necessarily represented statistically. The theory is based on the classical set theory, however with the modification that members belong to a particular set up to a certain degree within the interval [0, 1] [3.13]. The fuzzy function known as membership function is designed such that the value assigned to members within the fuzzy set denotes the level of the membership. Fuzzy set can be used when representing vagueness (fuzziness) and non-specificity in an uncertain variable [3.2, 3.13, 3.14]. The fuzzy set theory can be viewed more as a qualitative analysis tool rather than quantitative [3.10].

For a universal set *X*, the fuzzy membership function, μ_X obeys the following properties [3.11-3.14];

$$\begin{aligned} &\mu_{U-X}(x) = 1 - \mu_X(x) \text{ for } x \in U \\ &\mu_{X \cap Y}(x) = \min(\mu_X(x), \mu_Y(x)) \text{ for any } \quad x \in U \\ &\mu_{X \cup Y}(x) = \max(\mu_X(x), \mu_Y(x)) \text{ for any } \quad x \in U \end{aligned} \tag{3.2}$$

Members of the fuzzy set take values between 0 and 1 although the sum of the membership value does not always add up to unity like with probability [3.13]. The fuzzy set gives the possibility distribution function and not the probability distribution function of the output. A detailed explanation of the fuzzy set theory is presented in [3.13, 3.14].

3.2.3. Probability Theory

The probability theory is the most common approach of representing uncertainty especially when the distribution of the variables is available [3.15]. It provides a way of making quantitative inferences about uncertainties [3.10]. With the probability theory, the degree of likelihood or occurrence of a particular event is assigned a value known as probability. Mathematically, this is represented within a probability space [3.6].

A probability space consist of a set *S* called the space sample, an algebra χ over *S* whose elements are measureable and a probability measure $P : \chi \rightarrow [0,1]$. The probability theory is based on three axioms [3.15-3.17];

1. $P(x) \geq 0$ for all $x \in \chi$
2. $P(S) = 1$ (3.3)
3. for a disjoint set in χ $P(x_1, x_2, ...) = P(x_1) + P(x_2) + ...$

The probability measure is used when the uncertainty results from discordance or conflict among the likelihood of occurrence in elements of a set [3.14]. Thus, it is good at representing uncertainties due to stochastic disturbances, variability and risk [3.10]. Probabilistic analysis can be carried out if the probabilistic distribution of the uncertain variable is available. This is usually given in terms of the probability density function of the variable. This concept and other relevant probability terminologies are explained in the succeeding section.

To reflect the impact of the uncertain variable in the output, two stages are involved [3.2]; probabilistic representation of the uncertainty and propagation of the effect of the uncertainty through a model. In accounting for the effect of uncertainties in load flow studies, the probabilistic approach is commonly used since it is well established, precise and applicable to decision theory which is one of the objectives of the load flow study [3.10, 3.18]. One of the major drawbacks of the probabilistic approach is its high computational cost [3.18], however, this can be reduced with the proposition of approximate methods as used in this work.

3.3. Mathematical Basis

In this section, some basic terminologies and theories of probability, which are constantly used in the thesis, are discussed.

3.3.1. Definition of Terms

Random Experiment: An experiment whose outcomes are not known in advance. [3.19]

Random Event: An outcome or set of outcomes of a random experiment that have a definite probability of occurrence. [3.19]

Sample Space: A mathematical abstraction used to represent all possible outcomes of an experiment. [3.19, 3.20]

Random Variable: A random variable is a function that maps events defined on a sample space into a set of values. [3.19, 3.21]

Random Number: A number generated at a realization of any random variable that is an element of a given variate. [3.21]

Variate: A generalization of the idea of a random variable with similar probabilistic properties but defined without reference to a particular type of probabilistic experiment. A set of all random variables that obey a given probabilistic law [3.21].

3.3.2. Description of Random Variables

The probability that a variate **X** takes a value less or equal to a number *x* is termed the cumulative distribution function (CDF) *F(x).*

$$F(x) = \Pr[X \le x] \tag{3.4}$$

For the above to be true, *F(x)* must have the following properties [3.22];

i. Non-decreasing in *x* i.e. $F(x_2) \ge F(x_1)$ when $x_2 \ge x_1$

ii. Attain the maximum value of unity for the maximum *x*. $F(-\infty) = 0, F(\infty) = 1$

iii. Be continuous from the right

A distribution function can be continuous, discrete or mixed [3.21].

a. Continuous Distributions

A continuous distribution is characterized by *F(x)* being absolutely continuous. The continuous distribution has a monotonic non-decreasing nature as reflected by its properties [3.21, 3.22]. For a continuous distribution, (3.5) give the probability that *X* takes a value less or equal to *x*.

$$F(x) = \Pr[X \le x] = \int_{-\infty}^{x} f(u)du \tag{3.5}$$

where *f(x)* is the probability density function (PDF) of the random variable *X*.

The integral of the probability density function is always unity following (3.6).

$$\int_{-\infty}^{+\infty} f(x)dx = 1 \tag{3.6}$$

The PDF is the first derivate of the CDF of a random variable *X* and is given as;

$$f(x) = \frac{dF(x)}{dx} \tag{3.7}$$

Typical examples of continuous distributions include normal (Gaussian) distribution, Weibull distribution, and exponential distribution.

b. Discrete Distributions

A distribution is said to be discrete if *F(x)* has a fixed value except for a finite jump discontinuities [3.16]. The probability that *X* takes a value less or equal to *x* is given by;

$$F(x) = \Pr[X \leq x] = \sum_{x_i \leq x} p(x_i) \tag{3.8}$$

The binomial and Bernoulli distributions are examples of the discretely varying distribution. Discrete distributions also obey the properties listed above, hence, the sum of all the probabilities in a discrete distribution is one.

The mixed (composite) distribution is a distribution made of more than one continuous and or discrete distribution [3.21]. This is further discussed in section 3.3.5.

3.3.3. Moments and Cumulants of Random Variables

Moments of a random variable are the sums of its integral power that help to numerically describe the variable with respect to given characteristics such as location, variation, skew, peakedness etc. [3.23, 3.24]. Moments can be used in representing the nature of a random distribution. Although moments are commonly used, cumulants are alternative descriptive quantities of the random variable, which theoretically have properties superior to moments [3.24]. The cumulant (K_r) of order *r* of a variate can be calculated from all its moment not higher than order *r* and vice versa. The moment about the origin is referred to as raw moment (μ'_r) while the moment about the mean is called central moment (μ_r).

The raw and central moments for a continuous function *f(x)*, and a discrete distribution with p_k point probabilities are given in (3.9) and (3.10).

$$\mu'_r = \int_{-\infty}^{+\infty} x^r f(x)dx \qquad \text{(continuous)} \tag{3.9}$$

$$\mu'_r = \sum_{k=0}^{n} x_k^r p_k \qquad \text{(discrete)} \tag{3.10}$$

$$\mu_r = \int_{-\infty}^{+\infty} (x-\mu)^r f(x)dx \qquad \text{(continuous)} \tag{3.11}$$

$$\mu_r = \sum_{k=0}^{n} (x_k-\mu)^r p_k \qquad \text{(discrete)} \tag{3.12}$$

The first four moments are very popular as they give key information about the distribution. The first raw moment referred to as mean or expectation ($\overline{\mu}$), is a measure of central tendency. It gives the measure of tendency of the variables to cluster about the origin [3.23]. The second central moment called the variance (σ^2) measures the dispersion of the random variable about the mean point. The standard deviation (σ) of the variable is the square root of the variance. Another importance index is the coefficient of variation (CV) which is the ratio of the standard deviation to the mean.

$$CV = \frac{\sigma}{\overline{\mu}} \tag{3.13}$$

The degree of deviation of the random variable from the symmetry is measured using the normalised central third moment referred to as skewness (γ_1). The kurtosis (γ_2) which is the standardised central fourth moment gives the measure of the 'peakedness' of the random variable distribution [3.24].

$$\gamma_1 = \frac{\mu_3}{\sigma^3}, \gamma_2 = \frac{\mu_4}{\sigma^4} \tag{3.14}$$

a. Properties of Moments and Cumulants

i. Homogeneity.

All cumulants and moments (both raw and central) satisfy the homogeneity property [3.24]. For instance, if the variate *X* is multiplied by a constant, *a*, the moments and cumulants are in turn multiplied by a^r (*r* is the order of the moment/cumulant)

$$T_r(aX) = a^r T_r(X) \quad \text{where} \quad T = K, \mu, \mu' \tag{3.15}$$

ii. Invariantive Property

Cumulants (except the first (mean)) and central moments remain unchanged despite change in origin [3.24]. However, the first cumulant experiences a shift as illustrated below. The raw moments do not exhibit this property.

$$K_1(X + c) = K_1(X) + c \tag{3.16}$$

$$T_r(X + c) = T_r(X) \quad \text{where} \quad T = K, \mu \quad \text{and} \quad r \geq 2 \tag{3.17}$$

iii. Additivity

All cumulants (not moments) exhibit the additive property [3.24] that is for independent random variables X_1 and X_2, the cumulant is;

$$K_r(X_1 + X_2) = K_r(X_1) + K_r(X_2) \tag{3.18}$$

The additive property of cumulant is particularly useful in power systems. For instance, in determining the cumulant of injected power at a node, the cumulants of all the loads and generators on the bus can be easily added.

b. ***Relationship between Moments and Cumulants***

The relationship between moments and cumulants is given in (3.19) [3.24].

$$K_r = m_r - \sum_{j=1}^{r-1} \binom{r-1}{j-1} m_{r-j} K_j \qquad m = \mu \text{ or } \mu' \tag{3.19}$$

Using the above notation, the first 5 cumulant relative to the central and raw moments are shown in (3.20) and (3.21).

$$\begin{aligned} K_2 &= \mu_2 \ \text{(variance)} \\ K_3 &= \mu_3 \\ K_4 &= \mu_4 - 3\mu_2^{\,2} \\ K_5 &= \mu_5 - 10\mu_2\mu_3 \\ K_6 &= \mu_6 - 15\mu_4\mu_2 - 10\mu_3^{\,2} + 30\mu_2^{\,3} \end{aligned} \tag{3.20}$$

$$
\begin{aligned}
K_1 &= \mu_1' = \bar{\mu}\,(\text{mean}) \\
K_2 &= \mu'_2 - \mu_1'^2 \\
K_3 &= \mu'_3 - 3\mu_2'\,\mu_1' + 2\mu_1'^3 \\
K_4 &= \mu_4' - 4\mu_3'\,\mu_1' - 3\mu_2'^2 + 3\mu_2'\,\mu_1'^2 - 6\mu_1'^4 \\
K_5 &= \mu'_5 - 5\mu_4'\,\mu_1' - 10\mu'_3\,\mu_2' + 20\mu'_3\,\mu_1'^2 + 30\mu_2'^2\,\mu_1' - 60\mu_2'\,\mu_1'^3 + 24\mu_1'^5
\end{aligned}
\tag{3.21}
$$

3.3.4. Quantiles and Percentiles

Quantiles give a good idea of the general form of a distribution. For a random variable with ordered set of values, the quantile divides the probability into *q* equal parts. For a randomly distributed variable *x*, the k_{th} quantile is the value of x_k (in *x*) that corresponds to a cumulative frequency of *N.k/q* [3.25]. Depending on the value of *q*, quantiles can take different names. When the variable is divided into 100 equal parts (*q*=100), it is called percentile.

The quantile can also be estimated from the CDF of the random variable. The relationship between *q*-quantile and CDF is;

$$X(\mathrm{q}) = F^{-1}(q) \tag{3.22}$$

3.3.5. Mixed/Composite Distributions

A random variable is said to be a mixed distribution if it is a convex combination of other specific probability distribution functions [3.21]. The component distributions within the mixture can be finite or infinite. The mixture can also be composed of continuous distribution and or discrete distribution. A typical example of a mixed distribution comprising finite continuous component is a sample population made up of subpopulations [3.26] as illustrated in Fig. 3.1. The output wind power from a variable speed turbine is also a mixed distribution made up of finite continuous and discrete distributions.

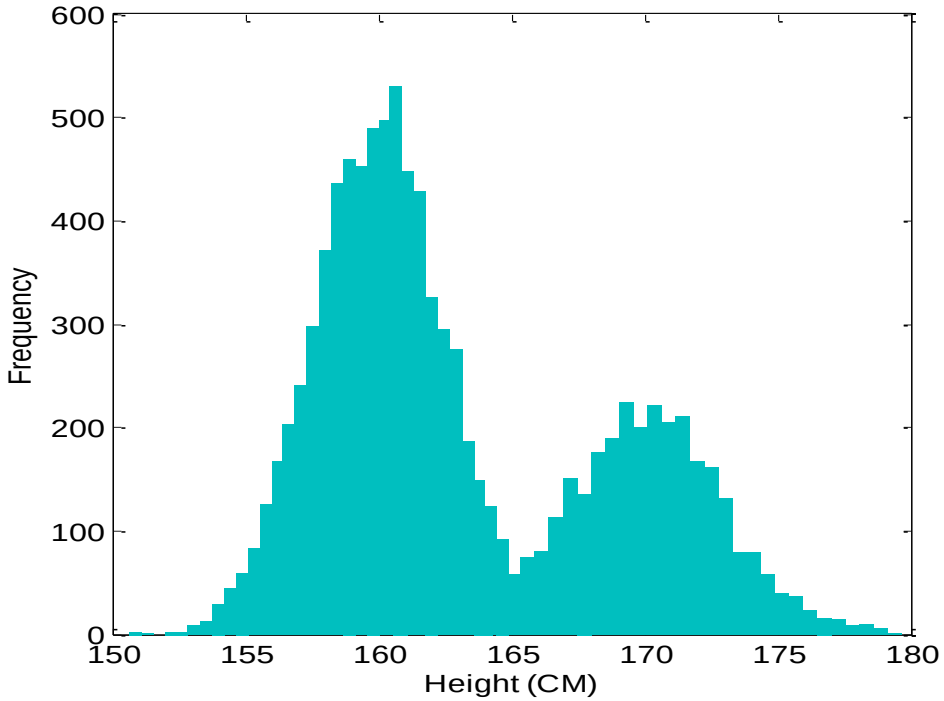

Fig 3.1: Histogram Showing a Mixed Distribution

The mixed distribution maintains the properties stated in section 3.3.2. The PDF and CDF of the mixed distribution are composed of the PDF and CDF of each constituent part allocated a weight to represent their significance/overall contribution to the mix. The mean and other moments of a mixed distribution are also derived from those of its composite distributions.

Let X be a mixture of two distributions $f_1(x)$ and $f_2(x)$, the final probability density function $f_c(x)$ can be represented as (3.23). Using this notation, the mean (μ'_c) is given as (3.24) [3.21, 3.26];

$$f_c(x) = p_1 f_1(x) + p_2 f_2(x) \tag{3.23}$$

$$\mu'_c = p_1\mu'_1 + p_2\mu'_2 \tag{3.24}$$

where p_1 and p_2 are the probability of having distributions $f_1(x)$ and $f_2(x)$ respectively and the sum of p_1 and p_2 must always be unity.

The n^{th} central moment μ_{cn} for a mixed distribution can be derived using (3.25), while substituting (3.23) in the equations.

$$\mu_{cn} = \int_{-\infty}^{\infty} (x - \mu'_c)^n f_c(x) dx \tag{3.25}$$

Expanding (3.23) and (3.25) above, the first five central moments of a mixed distribution having *i=1, 2, 3 … z* component distributions can be evaluated as below.

$$\begin{aligned}
\mu_{c2} &= \sigma_c^2 = \sum_{i=1}^{z} p_i[\mu_{i2} + (\mu'_i - \mu'_c)^2] \\
\mu_{c3} &= \sum_{i=1}^{z} p_i[\mu_{i3} + 3(\mu'_i - \mu'_c)\mu_{i2} + (\mu'_i - \mu'_c)^3] \\
\mu_{c4} &= \sum_{i=1}^{z} p_i[\mu_{i4} + 4(\mu'_i - \mu'_c)\mu_{i3} + 6(\mu'_i - \mu'_c)^2\mu_{i2} \ldots \\
&\quad \ldots + (\mu'_i - \mu'_c)^4] \\
\mu_{c5} &= \sum_{i=1}^{z} p_i[\mu_{i5} + 5(\mu'_i - \mu'_c)\mu_{i4} + 10(\mu'_i - \mu'_c)^2\mu_{i3} \ldots \\
&\quad \ldots + 10(\mu'_i - \mu'_c)^3\mu_{i2} + (\mu'_i - \mu'_c)^5]
\end{aligned} \tag{3.26}$$

3.3.6. Dependency

A set of variables are independent if the marginal probability associated with the outcome of one is not affected by the value observed for the other variable [3.21]. In line with the above definition, two variables will be dependent if and only if the probability of the occurrence of one is influenced by the other. Often times, dependence between variables is considered based on linear correlation between them. However, dependency goes beyond linearity as two variables with zero linear correlation can still be dependent. Some of the parameters for representing dependency include the Pearson product moment correlation (r), Spearman rank-order correlation coefficient (ρ), Kendall's Tau (τ) and Blomquist beta (β) [3.27].

It is also worth mentioning copulas which are functions that 'couple' or join marginal distributions to their joint distribution such that the dependence structure of the random variables can be captured [3.27, 3.28]. A detailed discussion on copulas is found in [3.27].

3.4. Approximate Techniques for Distribution Functions Reconstruction

The probability density function or the cumulative distribution function of a variable can easily be constructed if the moments or cumulants are known. Several techniques are available for this reconstruction with some of them based on series expansion of a base function [3.24]. In this section, the Gram Charlier series, the Cornish Fisher expansion series, the Edgeworth series and the Pearson system are briefly considered.

a. Gram-Charlier Type A Series Expansion

The Gram Charlier series expansion approximates the probability distribution of a variable in terms of the variables' cumulants [3.24, 3.29].

$$f(x) = \sum_{j=0}^{\infty} c_j H_j(x) \varphi(x) \tag{3.27}$$

where $\varphi(x)$ is the PDF of the standard normal distribution given by;

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \tag{3.28}$$

Multiplying (3.27) by the $H_r(x)$ (Chebyshev-Hermite polynomial) and integrating from -∞ to ∞ gives;

$$c_r = \frac{1}{r!} \int_{-\infty}^{\infty} f(x) H_r(x) dx \tag{3.29}$$

where $H_r(x)$ is built successively using (3.29) starting with H_o=1 [3.24].

$$H_r(x) = xH_{r-1}(x) - (r-1)H_{r-2}(x) \tag{3.30}$$

The resulting series in (3.31) is referred to as the Gram Charlier series of Type A [3.24, 3.29].

$$f(x)=\varphi(x)\left[\begin{array}{r} 1+\frac{K_3}{3!}H_3(x)+\frac{K_4}{4!}H_4(x)+\frac{10{K_3}^2}{6!}H_6(x)+\frac{K_5}{5!}H_5(x) \\ \cdots+\frac{35K_3K_4}{7!}H_7(x)+\frac{280{K_3}^3}{9!}H_9(x)+\cdots \end{array}\right] \tag{3.31}$$

where K is the cumulant.

$$g=\int_{-\infty}^{\infty} e^{\frac{x^2}{4}}\,dF(x) \tag{3.32}$$

The series can only be used if g (3.32), converges and if the PDF $f(x)$ of the variable tends to zero as |x| tends to infinity [3.30]. The main point in carrying out the approximation is to ensure the finite series give a good approximation of the original distribution [3.24]. The other issue has to do with the series giving negative values close to the tail in some instances [3.30]. The Gram Charlier series thus have a limited range of applicability and poor convergence [3.24] especially as $f(x)$ approaches infinity. This makes it inappropriate in practical problems where the tail of the distribution is of utmost importance.

b. Cornish Fisher Series Expansion

With the Cornish Fisher series, the quantiles of the function to be reconstructed are approximated using its cumulants and the quantiles of a standard normal distribution [3.30]. The cumulative distribution function is then estimated based on the inverse relationship between quantiles and cumulative distribution functions. The cumulative distribution function is then reconstructed using the first five cumulants of the random variable as shown in (3.33) [3.30]

$$\begin{aligned} x(q)\approx{}&\Phi_z^{-1}(q)+\frac{1}{6}(\Phi_z^{-1}(q)^2-1)K_3+\frac{1}{24}(\Phi_z^{-1}(q)^3-3\Phi_z^{-1}(q))K_4\ldots \\ &-\frac{1}{36}(2\Phi_z^{-1}(q)^3-5\Phi_z^{-1}(q))K_3{}^2+\frac{1}{120}(\Phi_z^{-1}(q)^4-6\Phi_z^{-1}(q)^2+3)K_5\ldots \\ &-\frac{1}{24}(\Phi_z^{-1}(q)^4-5\Phi_z^{-1}(q)^2+2)K_3K_4+\frac{1}{324}(12\Phi_z^{-1}(q)^4-53\Phi_z^{-1}(q)^2+7)K_3{}^3 \end{aligned} \tag{3.33}$$

where *x(q)* is the *q*-quantile function, $\Phi_z^{-1}(q)$ is the *q*-quantile of the standard normal distribution and κ_r is the r^{th} order cumulant of the distribution function *F(x).*

For non-Gaussian distributed random variables, the above procedure can be applied by standardising the variable such that its mean is zero and standard deviation unity. The higher moments are also standardised. In essence, the resulting *q*-quantiles are standardized and need to be transformed back to the original space of the distribution.

c. Edgeworth Series Expansion

The Edgeworth series is an inversion of the Cornish-fisher series used in reconstructing the probability density function of a variable [3.24]. The series employs the cumulants of the function and the Hermite polynomial as with the Gram Charlier series though, using a different arrangement and better degree of approximation [3.24]. The series is represented mathematically as shown in (3.34)

$$f(x) = \varphi(x) - \frac{K_3}{3!}\varphi^{(3)}(x) + \frac{K_4}{4!}\varphi^{(4)}(x) + \frac{10K_3^{\,2}}{6!}\varphi^{(6)}(x) - \frac{K_5}{5!}\varphi^{(5)}(x)\cdots - \frac{35K_3K_4}{7!}\varphi^{(7)}(x) - \frac{280K_3^{\,3}}{9!}\varphi^{(9)}(x)\cdots \qquad (3.34)$$

The Gram Charlier and the Edgeworth series have the same level of convergence, with both performing poorly for non-Gaussian distribution, although, the Edgeworth series performs better than the Gram Charlier series for distributions close to the Gaussian distribution [3.24, 3.30]. In addition, it has been shown to perform better for some non-Gaussian distributions as compared to the Gram Charlier series. In summary, the Edgeworth series can be seen to be fairly asymptotic thus giving values close to the true distribution [3.24, 3.31, 3.32]. The detailed treatment of the method and its performance comparison against other series is presented in [3.32].

d. Pearson System

The density function of a random variable can be reconstructed from its first four moments using the Pearson system [3.33]. For every member distribution of the Pearson system, the density function *f(x)* must satisfy (3.35) [3.24, 3.33-3.38].

$$\frac{df(x)}{dx} = \frac{(x-a)}{b_0 + b_1 x + b_2 x^2} f(x) \tag{3.35}$$

All members also obey the general result in (3.36).

$$x^n (b_0 + b_1 x + b_2 x^2) \frac{df(x)}{dx} = x^n (x-a) f(x) \tag{3.36}$$

Integrating the above equation from -∞ to ∞ gives (3.37).

$$-nb_0 \mu'_{n-1} - (n+1) b_1 \mu'_n - (n+2) b_2 \mu'_{n+1} = \mu'_{n+1} - a\mu'_n \tag{3.37}$$

where;

$$\mu'_n = \int_{-\infty}^{\infty} x^n f(x) dx \tag{3.38}$$

The constants *a*, b_0, b_1 and b_2 can be expressed in terms of the moments by solving (3.37) successively assuming *n* takes up values ranging from 0 to 3 and also taking the mean (μ'_1=0) as zero. The resulting values of the constants in terms of the moment are [3.24];

$$\left.\begin{aligned} a &= -\frac{\sqrt{\mu_2}\sqrt{\beta_1(\beta_2+3)}}{10\beta_2 - 12\beta_1 - 18} \\ b_0 &= -\frac{\mu_2(4\beta_2 - 3\beta_1)}{10\beta_2 - 12\beta_1 - 18} \\ b_1 &= a \\ b_2 &= -\frac{(2\beta_2 - 3\beta_1 - 6)}{10\beta_2 - 12\beta_1 - 18} \end{aligned}\right\} \tag{3.39}$$

β_1 and β_2 are respectively the square of the skewness (standardized third central moment) and the kurtosis (standardized fourth moment) while μ_2 is the second central moment (variance).

Depending on the distribution, there are seven main classifications based on the Pearson system [3.33]. The main decision as to which of the type fits a particular problem is based on using a criterion κ or from the plot of β_1 and β_2. The κ-criterion is estimated using (3.40) while the plot of β_1 and β_2 is shown in Fig. 3.2.

$$\kappa = \frac{b_1^{\,2}}{4b_0b_2} \tag{3.40}$$

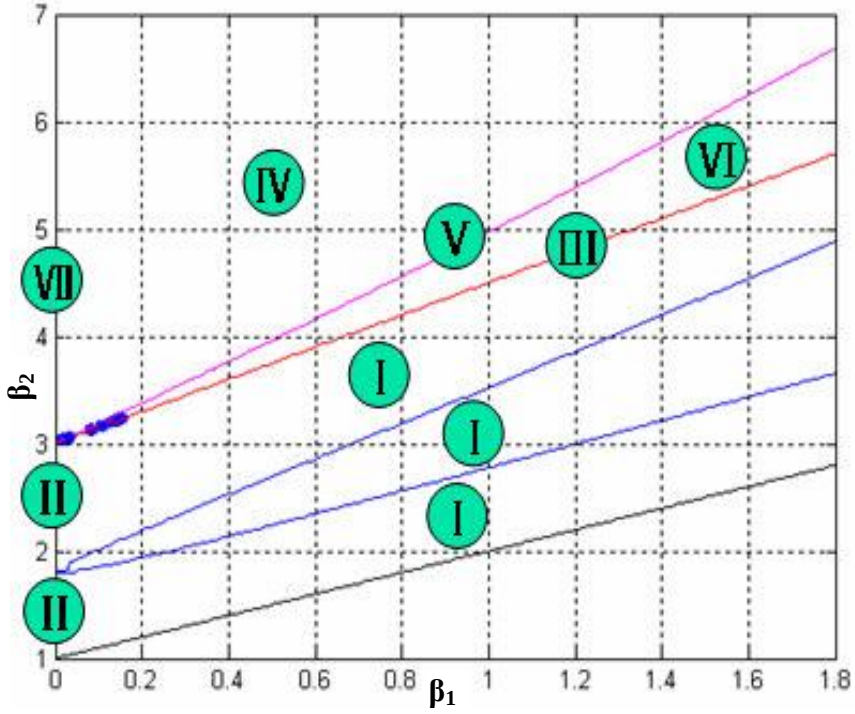

Fig 3.2: The Pearson Curve [3.33]

From the graph, the values of β_1 and β_2 give the Pearson distribution type to which the random variable fits. This subsequently determines the values of x and $f(x)$ in (3.35). The main limitation for the Pearson system is the restriction on the values of β_1 and β_2. To overcome this, transformation of the initial distribution into another space was suggested [3.37]. The other challenge is the singularity error when β_1 and β_2 for a particular random variable approach several distribution types [3.33]. A detailed treatment of the Pearson system including some restrictions and roots of each type are presented in [3.34-3.38].

3.5. Models of Uncertain (Random) Variables in Power Systems

As earlier identified, several uncertainties abound within the modern day power systems from generation to distribution. Some of the prominent ones include those due to the variability in consumer load demand and variability from both dispatchable and nondispatchable generators.

3.5.1 Important Distributions in Load Flow Studies

The statistical description of some frequently occurring distributions in load flow studies are discussed below.

a. Gaussian/Normal Distribution

A random variable x is normal or Gaussian if its density function is given by [3.16, 3.21];

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \tag{3.41}$$

where μ is the mean and σ^2 is the variance.

The cumulative distribution function of a normally distributed variable is;

$$F(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-(y-\mu)^2/2\sigma^2} dy \tag{3.42}$$

b. Bernoulli and Binomial Distributions

A random variable x is said to follow the Bernoulli distribution if it takes the value 0 or 1 while it is binomial distributed if it takes the value $0,1,2,\ldots,n$ for n parameters [3.16, 3.21]. The Bernoulli distribution is a special case of the binomial distribution. The distribution of the variable x is represented as;

$$P(x = k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad k = 0,1,2,\ldots,n \tag{3.43}$$

where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

The mean and variance are respectively

$$E(x) = np \tag{3.44}$$

$$\sigma^2 = np(1-p) \tag{3.45}$$

c. Weibull Distribution

A variable, which is Weibull distributed can either, follow the 2-parameter or 3-parameter Weibull distribution [3.39].

The density and distribution function for a variable following the 2-parameter Weibull is given in (3.46) and (3.47) respectively [3.39].

$$f(x) = \frac{\alpha}{\beta}(\frac{x}{\beta})^{\alpha-1}\exp[-(\frac{x}{\beta})^{\alpha}] \tag{3.46}$$

$$F(x) = 1-\exp[-(\frac{x}{\beta})^{\alpha}] \tag{3.47}$$

where α and β are respectively the shape and scale parameters.

The 3-parameter Weibull has an extra parameter x_o known as the location parameter and its density and distribution functions are given in (3.48) and (3.49) respectively.

$$f(x) = \frac{\alpha}{\beta}(\frac{x-x_o}{\beta})^{\alpha-1}\exp[-(\frac{x-x_o}{\beta})^{\alpha}] \tag{3.48}$$

$$F(x) = 1-\exp[-(\frac{x-x_o}{\beta})^{\alpha}] \tag{3.49}$$

3.5.1. Probabilistic Load Model

The probabilistic model of a load can be obtained by analysing series of measured data from the particular site. Depending on the load usage and nature, the model can follow a continuous distribution, or a discrete distribution and sometimes a combination of both. Due to lack of sufficient measured data (at a good interval), the Gaussian distribution is adopted throughout this work for the continuous load as with previous works. The discrete load follows the binomial or Bernoulli distributions. For instance, for a pumping machine, which is only able to work at five states, the load will have

certain probability of assuming any of the states. A typical discrete load distribution is illustrated in Table 3.1 [3.40].

Table 3.1: Typical Discrete Load Distribution

Active Power (MW)	13.4	19.6	30.2	34.8	37.3
Reactive Power (MVAR)	7.5	11.0	17.0	19.6	21.0
Probability	0.10	0.15	0.30	0.25	0.20

3.5.2. Probabilistic Model of A Generator

The dispatchable generator can either be dual state or multistate depending on the control mechanism put in place [3.40]. For a dual state conventional generator, full (operational) or null (outage) output is generated. The two-state generator follows the Bernoulli distribution while the multistate distribution follows the binomial distribution.

Most times, the randomness of the dispatchable generator output is expressed in form of its forced outage rate (FOR).

The "non-dispatchability" of the output of most renewable energy generators is dependent on the inability to fully control the supply of the fuel. The output from a solar PV farm is dependent on the solar insolation level while that for the wind farm is a function of the wind speed. The probabilistic model for the output solar power has been treated extensively in [3.41-3.43]. The model for the output wind power is summarized in the following section.

Other randomness within the network which can only assume two states such as the branch outage are represented using [0, 1] distribution [3.44].

3.5.3. Probabilistic Model of Wind Power

The probabilistic model of the output wind power from a turbine is dependent on the probabilistic model of the wind speed and the wind turbine characteristics curve. Hence, the wind speed probabilistic model needs to be

fully understood. In the past, wind speed has been represented probabilistically using the normal distribution, the gamma distribution and the Weibull distribution [3.30]. The wind speed distribution is dependent on the location and may vary in regions across the globe, although most researches have shown that the Weibull distribution is the best fit for the wind speed distribution for long term planning purposes [3.30]. To corroborate this, the wind speed data used in this work was obtained from real measurement for a location in Samsun, North Coast of Turkey.

The measured data for a 12-month period was analysed using a statistical software known as EasyFit [3.45]. Results obtained confirms the adequacy of the 2-parameter Weibull distribution in modelling the wind speed. A full analysis of the wind speed data is presented in Appendix B.

The variable wind speed generators (like the doubly fed induction generator, (DFIG)) are increasingly being used to ensure the power in the wind is properly harnessed at all times. As such, the standard mathematical relationship between wind speed and output wind power (3.50) will not always apply. The typical output from a variable speed wind turbine is thus employed as the basis for determining the probabilistic model of the output wind power.

$$P_w = \frac{1}{2}\rho A V^3 C_p \tag{3.50}$$

where P_w is the output wind power, ρ is the air density, A is the turbine blades swept area, V the wind speed and C_p is the power coefficient.

a. Approximate Model for wind speed-wind power relationship

The turbine characteristics curve (relationship between wind speed, output wind power and power coefficient) for a 500kW wind turbine [3.46] is shown in Fig. 3.3. For the purpose of analysis, a simplified form of the graph is shown in Fig. 3.4. From the literature, two models have been used in approximating this curve [3.47-3.53]. The first referred to as the linear model in this work is based on the assumption that region B in Fig 3.4 is linear. With this assumption, a simple linear relationship approximates the wind speed and wind power relationship. The second model will herewith be referred to as the cubic

model considers the initial relationship between wind speed and power (3.50) and thus assumes the wind power in region B is a cubic function of the wind speed.

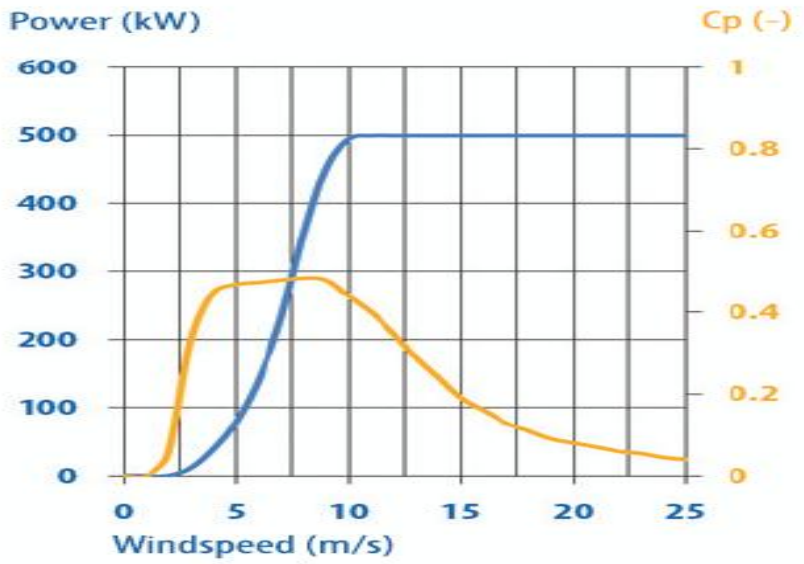

Fig 3.3: Manufacturer Wind Turbine Characteristics for EWT500 [3.46]

The adequacy of the two approximate models has been evaluated and results (presented in Appendix B) show that the cubic model gives a good representation of the real manufacturer's curve as opposed to the overestimation noticed with the linear model. The mathematical basis of the cubic model is further detailed below.

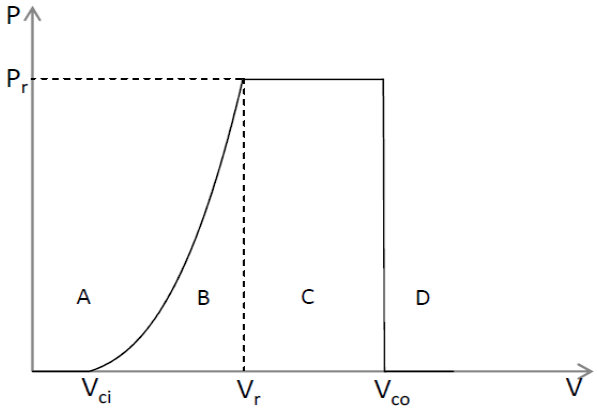

Fig 3.4: Power Output Vs. Wind Speed relationship for Wind Generators

The output wind power in each of these regions can be calculated using (3.51) for the cubic model [3.53].

$$P_w = \begin{cases} 0 & v \le v_{ci} \quad \text{(Region A)} \\ Cv^3 - D & v_{ci} < v \le v_r \quad \text{(Region B)} \\ P_r & v_r \le v \le v_{co} \quad \text{(Region C)} \\ 0 & v > v_{co} \quad \text{(Region D)} \end{cases} \tag{3.51}$$

where v_{ci} is the cut in wind speed, v_r is the wind speed at the rated power (P_r) and v_{co} is the cut out wind speed. Constant C and D are given as;

$$C = \frac{P_r}{(v_r^3 - v_{ci}^3)} \quad , \qquad D = Cv_{ci}^3 \tag{3.52}$$

Combining equations (3.51) and (3.52) with the wind speed distribution (3.47) result into (3.53) which is the cumulative distribution.

$$F(P_w) = \begin{cases} 1 + \exp\left[-\left(\frac{v_{co}}{\beta}\right)^{\alpha}\right] - \exp\left[-\left(\frac{v_{ci}}{\beta}\right)^{\alpha}\right] & P_w = 0 \\ 1 + \exp\left[-\left(\frac{v_{co}}{\beta}\right)^{\alpha}\right] - \exp\left[-\frac{((P_w + D)/C)^{\alpha/3}}{\beta^{\alpha}}\right] & 0 < P_w < P_r \\ 1 & P_w = P_r \end{cases} \tag{3.53}$$

Differentiating the CDF (3.53) with respect to output power P_w, the probability density function (PDF) is obtained as;

$$f(P_w) = \begin{cases} 1 + \exp\left[-\left(\frac{v_{co}}{\beta}\right)^{\alpha}\right] - \exp\left[-\left(\frac{v_{ci}}{\beta}\right)^{\alpha}\right] & P_w = 0 \\ \frac{\alpha}{3C\beta^{\alpha}}\left(\frac{P_w + D}{C}\right)^{(\alpha/3)-1} \exp\left[-\frac{((P_w + D)/C)^{\alpha/3}}{\beta^{\alpha}}\right] & 0 < P_w < P_r \\ \exp\left[-\left(\frac{v_r}{\beta}\right)^{\alpha}\right] - \exp\left[-\left(\frac{v_{co}}{\beta}\right)^{\alpha}\right] & P_w = P_r \end{cases} \tag{3.54}$$

It is worth stating that in some previous literature [3.47-3.52] on probabilistic load flow using approximate methods, the probabilistic model of the wind speed is used directly in determining evaluation points for the load flow, however, this technique introduces some errors as explained in Appendix B.

b. Output Wind Power as a Composite Distribution

Using the wind power model described above, the output from a 20MW wind farm is show in Fig. 3.5. From the plot, the output wind power is seen to compose of three distinct parts; a discrete component at zero power part, the continuous part reflecting Region B in Fig. 3.4 and another discrete part at the rated power.

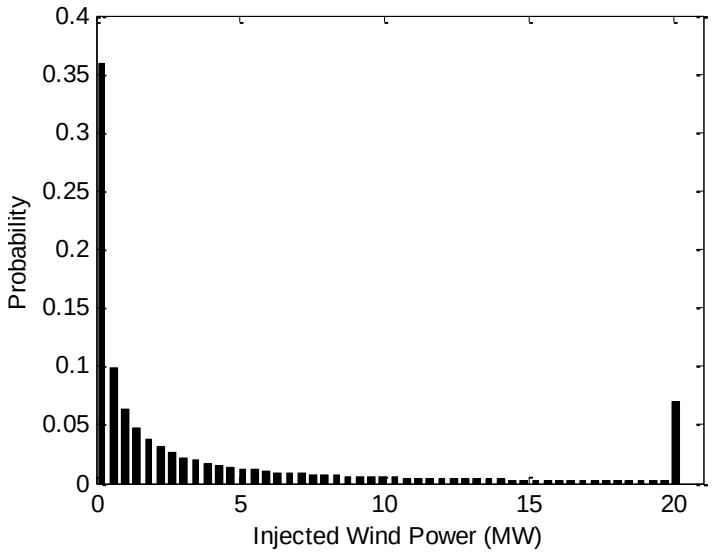

Fig 3.5: Injected Wind Power from a 20MW Wind Farm

To drive further the mixed distribution attribute of wind power, Fig. 3.6 shows the typical output distribution of a 1MW wind turbine installed in a high wind profile region with cut-in, rated, and cut-out wind speeds of 4m/s, 15m/s and 25m/s respectively [3.49]. The wind speed parameter based on the 2-parameter Weibull distribution are α= 3.97 (shape parameter) and β=10.7 (scale parameter) [3.49]. As expected, the PDF of the output wind power is continuous in the region where $P_w<P_r$ while the distribution becomes truncated and remains constant at the rated power after this point.

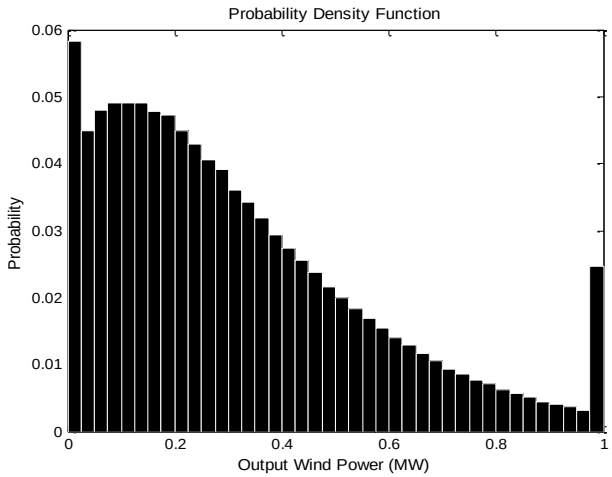

Fig 3.6: Typical wind power output from a 1MW Turbine.

3.6. Précis

Power systems like most physical systems are prone to uncertainties from various sources ranging from natural (uncontrollable) to those as a result of imperfection in system and measurement. Based on the type and source of the uncertainty, several classification of uncertainty can be found in the literature. The common classification into aleatory and epistemic uncertainty has been discussed in this chapter. Aleatory uncertainties, mostly due to natural variability, are seen to be irreducible while epistemic uncertainties can generally be reduced by employing the "right" or more sophisticated model, method or process. Uncertainties due to output wind power are typical aleatory uncertainties since it is dependent on the weather condition which is uncontrollable while uncertainties due to the utilisation of various probabilistic load flow models can be viewed as epistemic since it is reducible through the use of an appropriate method.

The various ways on representing uncertainties have been highlighted while a brief discussion of interval mathematics, fuzzy set theory and probabilistic theory previously used in representing uncertainties in power systems are given. Considering the merits of probabilistic load flow in decision based application like load flow studies, an in-depth explanation including the mathematical basis of the concept is presented.

To accurately reflect the effect of uncertainties once they are projected through the model, the statistic and distribution and density plot of the output is desired. Four systems of approximating the density and or distribution plot namely the Gram-Chalier series, Edgeworth series, Cornish-Fisher series and the Pearson System have been examined. The mathematical basis for each system has been presented alongside their limitations.

In representing uncertainties, the mathematical model/distribution of the random variable must be known. This is often obtained by studying the source, nature and or effect of the variable. For instance, in power systems, the model of uncertainties due to the customer load demand is obtained by analysing the recorded data over a period of time. The probabilistic models of other uncertainties due to both dispatchable and non-dispatchable generators are discussed in this chapter. Considering that wind power is one of the most viable of the non-dispatchable (renewable) power sources, its probabilistic model is of interest. Two models (linear and cubic) have been previously employed in modelling output wind power. The accuracy of the models has been examined in this chapter. A validation of the models was carried out by comparing results obtained using them with those from real wind speed measurement plus the manufacturer's curve. From the results, the cubic model was seen to be the most adequate since the linear model overestimates wind power.

3.7. References

[3.1] W. L. Oberkampf, J. C. Helton, C. A. Joslyn, S. F. Wojtkiewicz, and S. Ferson, "Challenge Problems: Uncertainty in System Response Given Uncertain Parameters," [Online]. Available: http://www.sandia.gov/epistemic/prob.statement.12-01.pdf [accessed 28/11/2012]

[3.2] S.S. Isukapalli, "Uncertainty Analysis of Transport-Transformation Models", PhD. Thesis, The State University of New Jersey. 1999

[3.3] Uncertainty and error, [Online]. Available: http://www.columbia.edu/cu/physics/pdf-files/Lab_1-01.pdf [accessed 29/11/2012]

[3.4] D. B. Pengra and L. T. Dillman "Notes on Data Analysis and Experimental Uncertainty," [Online]. Available: http://courses.washington.edu/phys431/uncertainty_notes.pdf [accessed 29/11/2012]

[3.5] T. Rosqvist, "Types of uncertainties in the modelling of complex systems" Presentation at the International Seminar on Climate Change impact assessment and adaptation, [Online]. Available: http://www.vatt.fi/file/tolerate_may08_tro.pdf

[3.6] A. Ajayi, "Direct Computation of Statistical Variations in Electromagnetic Problems," PhD Thesis, University of Nottingham, May 2008.

[3.7] "Interval Analysis" *Encyclopedia of Mathematics* [Online]. Available: http://www.encyclopediaofmath.org/index.php/Interval_analysis [accessed 05/12/2012]

[3.8] J.G. Rokne, "Interval Arithmetic and Interval Analysis: An Introduction," [Online]. Available: http://pages.cpsc.ucalgary.ca/~rokne/CPSC491/interval_art.pdf [accessed 27/11/2012]

[3.9] H. Dawood, *Theories of Interval Arithmetic*, [Online], LAP Lambert Academic Publishing GmbH & Co, 2011

[3.10] F. Alvarado, Y. Hu, and R. Adapa "Uncertainty in Power System Modeling and Computation," In *proc. IEEE International Conference on Systems, Man and Cybernetics*, pp 754-760, Oct. 1992

[3.11] R. R. Yager, "Uncertainty Representation using Fuzzy Measures," *IEEE Trans. on Systems man and Cybernetics*, vol. 32, no. 1, pp13-20, Feb 2002.

[3.12] Z. Pawlak, "Rough Sets", [Online]. Available: http://bcpw.bg.pw.edu.pl/Content/2026/RoughSetsRep29.pdf

[3.13] G. J. Klir and T. A. Folger, *Fuzzy Sets, Uncertainty and Information*, New Jersey: Prentice Hall, 1988

[3.14] G. J. Klir, *Uncertainty and Information: Foundation of Generalized Information theory* John Wiley & Sons, Inc. 2006

[3.15] M. A. Bean, *Probability: The Science of Uncertainty*, Pacific Grove, CA, 2001

[3.16] A. Papoulis and S. U. Pillai, "Probability, Random Variables and Stochastic Processes", McGraw-Hill, 4th edition 2002

[3.17] R. Fagin and J. Y. Halper, "Uncertainty, belief and probability", *Journal of Computational Intelligence*. vol. 7, pp160-172, 1991

[3.18] J. Halpern, "Plausibility Measures: A general Approach for Representing Uncertainty," [Online]. Available: http://www.uni-konstanz.de/philosophie/fe/files/halpern_plausibility_measures.pdf [accessed 26/11/2012]

[3.19] W. Feller, *An introduction to probability theory and its applications*, John Wiley & Sons, 1968

[3.20] V.V. Rao, "Probability and Random Variables" [Online]. Available: http://nptel.iitm.ac.in/courses/IIT-MADRAS/Principles_of_Communication1/Pdfs/1_5.pdf [accessed 23/11/2012]

[3.21] B. Forbes, M. Evans, N. Hastings and B. Peacock, *Statistical Distributions*, 4th ed., New Jersey: John Wiley, 2011.

[3.22] M. Abramowitz, and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, New York: Dover, pp. 927-961, 1972.

[3.23] W. H. Press S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, *Numerical Recipes in C*, 2nd ed., New York: Cambridge University Press, 1997, pp 610-613.

[3.24] A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics*, 6th ed., vol. 1. London: Edward Arnold, 1994.

[3.25] E. W. Weisstein, "Quantile," *MathWorld* A Wolfram Web Resource, [Online]. Available: http://mathworld.wolfram.com/Quantile.html

[3.26] M. H. Al-Haboubi, "Statistics for a composite distribution in anthropometric studies" *Journal of Ergonomics*, vol.40, no.2, pp.189-198, 1997

[3.27] M. S. Khadka, J. Y. Shin, N. J. Park, K.M. George, and N. Park, "Copula-Based Density Weighting Functions", [Online]. Available: www.cs.okstate.edu/ferl/documents/increse06.pdf, accessed Dec 2011

[3.28] R. B. Nelsen, *Introduction to Copulas*, Springer , second edition 2006

[3.29] Gram-Charlier Series. *Encyclopedia of Mathematics,* [Online]. Available: http://www.encyclopediaofmath.org/index.php/Gram-Charlier_series [accessed 27/11/2012]

[3.30] J. Usaola, "Probabilistic load flow with wind production uncertainty using cumulants and Cornish fisher expansion," *International Journal of Electerical Power and Energy Systems*, vol. 31, iss 9, pp. 474-481, Oct. 2009.

[3.31] Edgeworth Series. *Encyclopedia of Mathematics,* [Online]. Available: http://www.encyclopediaofmath.org/index.php/Edgeworth_series [accessed 27/11/2012]

[3.32] M. Fan, V. Vittal, G. T. Heydt, and R. Ayyanar, "Probabilistic Power Flow Studies for Transmission Systems With Photovoltaic Generation Using Cumulants" *IEEE Trans. on Power Systems*, vol. 27, no. 4, pp. 2251-2261, Nov., 2012.

[3.33] B. D. Youn, Z. Xi, L. J. Wells, and P. Wang, "Enhanced Dimension-Reduction (eDR) Method for Sensitivity-Free Uncertainty Quantification," [Online]. Available: http://www.me.mtu.edu/~bdyoun/files/eDR.pdf [accessed 27/11/2012]

[3.34] "Alignment of Statistical Distributions. Pearson's curves" in Applied & Engineering Mathematics [Online]. Available: http://www.simumath.com/library/contents.html [accessed 23/11/2012]

[3.35] M. A. Srokosz "A New Statistical Distribution for the Surface Elevation of Weakly Nonlinear Water Waves," *American Meteorological Society Journal of Physical Oceanography* vol. 28, pp. 149-155

[3.36] A. Andreev, A. Kanto and P. Malo "Computational Examples of a New Method for Distribution Selection in the Pearson System," *Journal of Applied Statistics*, vol. 34, no.4, 487-506, May 2007

[3.37] H. Solomon and M. A. Stephens "Approximations to Density Functions Using Pearson Curves," *Journal of the American Statistical Association*, vol. 73, no. 361, pp. 153-160, Mar., 1978.

[3.38] N. L. Johnson, "Systems of Frequency Curves Generated by Methods of Translation," Biometrika, vol. 36, no. 1/2, pp. 149-176, Jun., 1949.

[3.39] H. Rinne, *The Weibull distribution*, Taylor & Francis, 2009

[3.40] X. Wang, J. R. McDonald, *Modern Power Systems Planning*, McGraw-Hill 1994.

[3.41] S. Conti and S. Raiti, "Probabilistic Load Flow for Distribution Networks with Photovoltaic Generators Part 1: Theoretical Concepts and Models," In *proc. 2007 IEEE International Conference on Clean Electric Power*, pp. 132-136

[3.42] S. Conti, S. Raiti, and C. D. Gregorio "Probabilistic Load Flow for Distribution Networks with Photovoltaic Generators Part 2: Application to a Case Study," In *proc. 2007 IEEE International Conference on Clean Electric Power*, pp137-141

[3.43] F.J. Ruiz-Rodriguez, J.C. Hernandez, and F. Jurado, "Probabilistic load flow for radial distribution networks with photovoltaic generators," *IET Renewable Power Generation*, vol.6, Iss. 2, pp.110-121, 2012.

[3.44] Z. Hu and X. Wang, "A probabilistic load flow method considering branch outages," *IEEE Trans., Power. Systems* vol.21, no. 2, pp. 507-514, May 2006.

[3.45] EasyFit, ©MathWave Technologies, [online]. Available: http://www.mathwave.com/products/easyfit.html

[3.46] EWT500, [online]. Available: http://www.genatec.co.uk/downloads/EWT.pdf [accessed 11/01/2012]

[3.47] D. Lei, C. Weidong, B. Hao and Y. Yi-han, "Probabilistic load flow analysis for power system containing wind farms," in *Proc. 2010 APPEEC Power and Energy Engineering Conf.*

[3.48] T. Gui-Kun, Y. Shu-Jun, and W. Yan, "Research on Probabilistic Power Flow of the Distribution System Based on Cornish-Fisher," *in Proc. 2011, Power and Energy Engineering Conference (APPEEC), Asia-Pacific.*

[3.49] Z. Bie, G. Li, H. Liu, X. Wang and X. Wang, "Studies on Voltage Fluctuation in the Integration of Wind Power Plants Using Probabilistic Load Flow," *presented at the Power and Energy Society General meeting, Pittsburgh.* 20-24 July 2008

[3.50] D. M. Outcalt. "Probabilistic Load Flow for High Wind Penetrated Power Systems based on a Five Point Estimation Method," PhD Thesis, University of Wisconsin-Milwaukee, Dec. 2009.

[3.51] O. A. Oke, D. W. P. Thomas, G. M. Asher and L. R. A. X. de Menezes, "Probabilistic Load Flow for Distribution Systems with Wind Production using the Unscented Transforms Method," *In pro. Innovative Smart Grid Technologies (ISGT), 2011 IEEE PES*

[3.52] O. A. Oke, D. W. P. Thomas and G. M. Asher, "A New Probabilistic Load Flow Method for Systems with Wind Penetration,*" in Pro. PowerTech 2011 Trondheim,* 19-23 June 2011

[3.53] H. Bayem, M. Petit, Ph. Dessante, F. Dufourd and R. Belhomme, "Probabilistic characterization of wind farms for grid connection studies," *European Wind Energy Conference & Exhibition*, May 7-10 2007, Milan, Italy.

Chapter 4

Probabilistic Load Flow Methods

Probabilistic load flow studies are pivotal in determining the true power system state in the presence of uncertainties. The effectiveness of any probabilistic load flow method can be measured in terms of execution time, accuracy, as well as simplicity. This chapter examines in detail some previously proposed probabilistic load flow methods in the literature while also demonstrating their working principles.

4.1. Method Classifications

Probabilistic load flow (PLF) methods have been broadly classified based on the mathematical techniques on which they operate. These methods can be divided into; numerical or simulation methods [4.1-4.6], analytical methods [4.7-4.21], approximate methods [4.23-4.31] and hybrid methods [4.32-4.34]. The Fuzzy load flow technique [4.35-4.38] and Interval arithmetic load flow [4.39, 4.41] methods which present alternative ways of representing the uncertainty are also worthy of mention. However, the most prominently used methods fall under the first three classifications. Example techniques under each of these are further discussed.

The non-linear load flow equation to be solved can be represented by (4.1 and 4.2) [4.17].

$$\left.\begin{aligned} P_i &= V_i \sum_{k=1}^{n} V_k (G_{ik} \cos\theta_{ik} + B_{ik} \sin\theta_{ik}) \\ Q_i &= V_i \sum_{k=1}^{n} V_k (G_{ik} \sin\theta_{ik} - B_{ik} \cos\theta_{ik}) \end{aligned}\right\} \tag{4.1}$$

$$\left.\begin{aligned} P_{ik} &= V_i V_k (G_{ik} \cos\theta_{ik} + B_{ik} \sin\theta_{ik}) - G_{ik} (t_{ik} V_i)^2 \\ Q_{ik} &= V_i V_k (G_{ik} \sin\theta_{ik} - B_{ik} \cos\theta_{ik}) + (B_{ik} - b_{ik}) V_i^2 \end{aligned}\right\} \tag{4.2}$$

where P_i and Q_i are the net active and reactive powers injected at node i and θ_{ik} is the voltage phase angle between nodes i and k. P_{ik} and Q_{ik} are the active and reactive power flows in line ik, V_i and V_k are respectively the voltage magnitude at bus i and k, G_{ik} and B_{ik} are the real and imaginary parts of the network admittance matrix while t_{ik} is the transformation ratio of branch ik. For a branch with transmission line t_{ik} is unity. b_{ik} is half the branch susceptance for transmission line, while it is neglected for branches with in-phase or phase shift transformers.

4.2. Numerical or Simulation Methods for PLF studies

Numerical methods are based on producing a sequence of approximations by repeatedly solving the problem [4.42]. Monte Carlo Simulation (MCS) is a typical numerical method [4.2-4.3] often used in probabilistic load flow studies.

Monte Carlo simulation techniques are based on sampling the variables at random, following the probability density function (PDF) of the stochastic variables and then observing the results [4.43]. In using the MCS method, the distribution function of all the uncertain variables within the power network should be known. The MCS method comprises of two vital features [4.1]; the random number generation and the random sampling. The random regenerator is easily built using some basic mathematical techniques once the PDF is given. While several sampling techniques exist, the commonly used technique in PLF studies is the simple random sampling (SRS) because of its simplicity. This is often referred to as the crude MCS.

Using the crude MCS method, the non-linear deterministic load flow equations (4.1) are sequentially solved using the randomly generated variables. The key advantage of the MCS method is its ability to use the exact non-linear equations thus eliminating linearization errors [4.1]. However,one major drawback of the crude MCS is that the number of samples has to increase for the accuracy to increase [4.43-4.46]. This implies that a large number of samples have to be generated before it converges. This makes the crude MCS method computationally burdensome and it also requires a large computer storage space.

To further illustrate home the significant relationship between sample numbers and convergence rate, a simple 6 bus test system [4.48] is considered. The system is assumed to have three normally distributed randomly varying functions with a 5% coefficient of variation. A graph showing the convergence rate and time against the number of simulation samples is presented in Fig. 4.1.

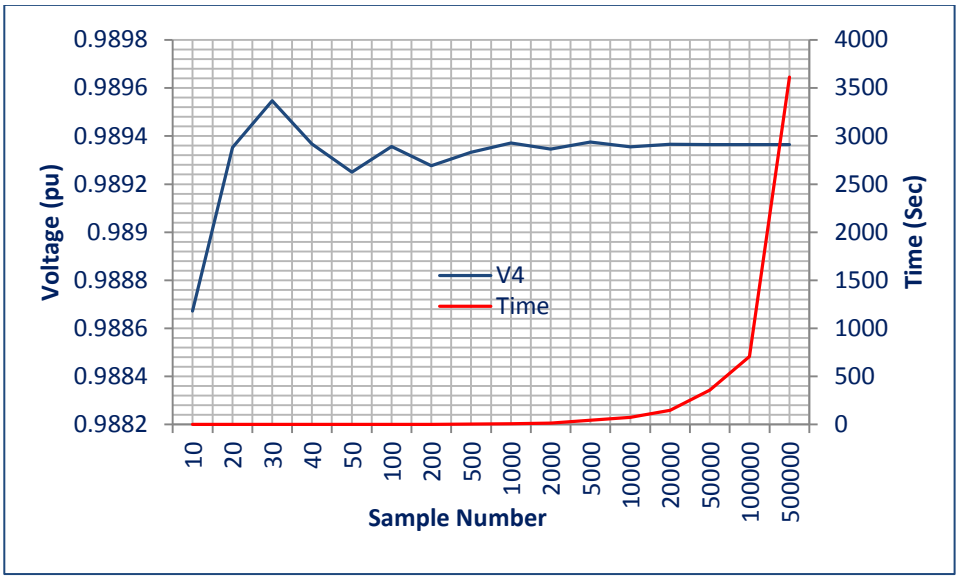

Fig 4. 1: Convergence rate and computation time against the sample size using the MCS method

From the graph, it is clear that a large sample number is required before convergence can be attained. For instance, a minimum of 20000 samples with simulation time of about 150 seconds was required in this case before convergence was attained. Although the crude MCS gives accurate results (when a large sample size is used) and often used as the benchmark in evaluating the performance of other methods, its prohibitive computational time especially for large systems makes it unattractive.

To reduce the sample size using the crude MCS method, alternative techniques such as the Latin Hypercube Sampling (LHS) [4.43, 4.46 4.47] and various reductions of variance techniques [4.43, 4.45-4.47] can be used in tackling the sample size problem. Examples of variance reduction techniques include importance sampling, stratified sampling, correlation and regression, mixed sampling, control variates, antithetic variates etc. Generally, variance reduction techniques sample the variable in 'interesting' areas based on *a priori*

information [4.45], in the real sense, the more information available, the better the accuracy that is achievable [4.43]. To have this knowledge, it is advised that the simple MCS can be used for the initial simulation, while these methods can be applied for the subsequent simulations [4.43]. The other main drawback of these techniques is with the modifications they introduce to the probability distribution of the variates [4.46].

Of the various improvement techniques, the Latin Hypercube Sampling (LHS) [4.4-4.6] seems to have gained prominence in power systems applications. The LHS is a combination of stratified sampling and random sampling which samples the variate over the entire distribution. With the LHS method, the entire parameter space is divided into segments having the same probability, samples are then drawn from each segment randomly or by some efficient means such as using the midpoint [4.46]. Cells (on the same row and column) having the sample value as the chosen cell are eliminated [4.46]. The process is repeated until the final samples are obtained. One challenge with this technique has to do with its elimination of the possible interaction between variables [4.46]. However, the combined technique used in [4.4]-[4.6] suggests that this dependence problem can be reduced.

Fewer number of samples are required using the LHS as compared with the simple MCS, however, for a multivariate problem, samples for all the variables have to be generated at once thus leading to a need for a large computer memory storage space [4.44]. Also, the technique only performs well for the evaluation of the expectation but may be unsuitable if the distribution of the output variables is desired [4.4]. These drawbacks led to the search for alternatives that are able to incorporate the advantages of the MCS but void of its setbacks.

4.3. Analytical Methods

Analytical methods were proposed as alternatives to the MCS method in view of their computational time disadvantage. Analytical methods can be referred to as mathematical techniques which involve the operation of the probability density function (PDF) of the input variables (power) so as to obtain the PDFs

of the output variables (system state and line flows) [4.11]. Various techniques have been previously used in carrying out this transformation, though, all analytical based techniques require some form of assumptions before they can be implemented. Some of these assumptions include [4.1, 4.10];

i. Linearization of the load flow equation given in (4.1) and (4.2)
ii. Constant structure of the network
iii. Independence between input parameters

The initial techniques proposed [4.7] involve the convolution of the density functions of the input variable to obtain the density function of the outputs. Two techniques were proposed for carrying out the convolution process, the first often referred to as the conventional method is based on using Laplace transforms [4.8] and another technique based on the fast Fourier transforms (FFT) [4.9-4.11]. Unfortunately, using both methods involve complicated mathematics [4.1], result in large errors due to linearization, while errors also tend to increase as the variation in the input parameters increases. Improvements such as multipoint linearization [4.11], network outage modelling [4.10] and dependence representation [4.12] have been proposed for the conventional and FFT methods, although the underlying problem caused by linearization still poses a great problem to their wide application.

Recently, another analytical method known as the Cumulant Method (CM) has gained attention due to its computational speed advantage [4.13]. This method is further discussed below.

4.3.1. The Cumulant Method for PLF Studies

The Cumulant method exploits the unique properties of cumulants as discussed in section 3.3.3, thus avoiding the complicated convolution of PDF's as with the conventional and FFT based analytical methods. The distribution of the desired output parameters are evaluated from those of the input parameters just like with other analytical methods using some key relationship extracted from the linearized equations. The load flow equations given in (4.1) and (4.2) are linearized to give [4.17];

$$\left.\begin{aligned}\mathbf{W} &= f(\mathbf{X})\\ \mathbf{Z} &= g(\mathbf{X})\end{aligned}\right\} \tag{4.3}$$

where ***W*** is the nodal active and reactive power injection vector at the normal operating conditions, ***Z*** is the branch flow vector under normal operating conditions, ***X*** is the state vector comprising of the nodal voltages and angles under normal operating conditions while *f* and *g* are respectively the power injection and line flow functions.

At the normal operating conditions, (4.3) becomes (4.4).

$$\left.\begin{aligned}\mathbf{W}_o &= f(\mathbf{X}_o)\\ \mathbf{Z}_o &= g(\mathbf{X}_o)\end{aligned}\right\} \tag{4.4}$$

Assuming the injected power in the system is subjected to a small disturbance $\Delta \boldsymbol{W}$, there will be a change $\Delta \boldsymbol{X}$ and $\Delta \boldsymbol{Z}$ in the state variable and branch flow vector respectively. If the disturbance is not very large, expanding (4.3) around the operating point using Taylor series gives;

$$\mathbf{W} = \Delta\mathbf{W} + \mathbf{W}_o = f(\mathbf{X}_o) + f'(\mathbf{X}_o)\Delta\mathbf{X} + \cdots \tag{4.5}$$

$$\mathbf{Z} = \Delta\mathbf{Z} + \mathbf{Z}_o = g(\mathbf{X}_o) + g'(\mathbf{X}_o)\Delta\mathbf{X} + \cdots \tag{4.6}$$

Ignoring the higher order terms, (4.5) and (4.6) become;

$$\Delta\mathbf{W} = f'(\mathbf{X}_o)\Delta\mathbf{X} + \cdots \tag{4.7}$$

$$\Delta\mathbf{Z} = g'(\mathbf{X}_o)\Delta\mathbf{X} + \cdots \tag{4.8}$$

where

$$f'(\mathbf{X}_o) = \left.\frac{\delta f(\mathbf{X})}{\delta \mathbf{X}}\right|_{\mathbf{X}=\mathbf{X}o} = \mathbf{J}_o \tag{4.9}$$

$$g'(\mathbf{X}_o) = \left.\frac{\delta g(\mathbf{X})}{\delta \mathbf{X}}\right|_{\mathbf{X}=\mathbf{X}o} = G_o \tag{4.10}$$

The change in the state variable $\Delta \boldsymbol{X}$**,** is obtained by merging (4.7) and (4.9). where $\boldsymbol{J}_o$ is the Jacobian matrix and $\boldsymbol{S}_o$ is the sensitivity matrix.

$$\Delta \mathbf{X} = \mathbf{J}_o^{-1} \Delta \mathbf{W} = \mathbf{S}_o \Delta \mathbf{W} \tag{4.11}$$

The change in the power flow in the branches $\Delta \mathbf{Z}$ is given as [4.17];

$$\Delta \mathbf{Z} = \mathbf{G}_o \Delta \mathbf{X} \tag{4.12}$$

Substituting (4.11) into (4.12) gives;

$$\Delta \mathbf{Z} = \mathbf{G}_o \mathbf{S}_o \Delta \mathbf{W} \tag{4.13}$$

With the above equations, any change in the injected power is easily transformed to the outputs (state variables and branch flows) once the cumulants of the injected power are known. The Cumulant method is often combined with one of Cornish-Fisher Approximate series [4.15, 4.18, 4.19], Gram Charlier Series [4.13, 4.16, 4.17-4.19] or the Egdeworth expansion [4.18, 4.19] in getting the distribution of the output variables. A performance comparison of these distribution reconstruction schemes is given in [4.15, 4.19]. Although most of the propositions [4.13-4.19] ignored the effect of dependency between variables, an orthogonalization technique is used in [4.20] to include the linear correlation effect in the cumulant of the variables. A detailed treatment of the cumulant method is found in [4.17].

The additive properties of cumulants (section 3.3.3) helps avoid some of the complicated mathematics associated with other analytical methods, however, errors due to linearization still make the method inadequate especially when the input distribution have some active points aside from the mean value [4.20]. Some enhancements [4.20, 4.21], have been proposed which are all based on solving the load flow at other active points aside from the expected value. These enhancements give better level of accuracy as compared with the original CM, though, the results (especially for higher moments) are still prone to significant errors due to its dependence on the Jacobian matrix [4.22].

4.4. Approximate Methods for PLF Studies

The general aim of any approximate method is to represent non-arithmetic quantities with arithmetic quantities to a good level of accuracy using minimal

computation time and effort [4.49]. Approximate methods such as First order second moment method (FOSMM) [4.23, 4.24], point estimate method (PEM) [4.25-4.30] and more recently the unscented transform method (UT) [4.31] have been used in solving the probabilistic load flow problem. These approximate methods have gradually gained prominence in carrying out probabilistic load flow due to their robust ability to directly make use of the non-linear load flow equations like the MCS method, thus, eliminating the linearization errors associated with the analytical methods.

The FOSMM was first used in [4.23] for optimal load flow studies. The method involves using the first moment (mean) of the input function in obtaining the first and second moments of the output function by using the Taylor series expansion in approximating the non-linear equation. The method used here relates closely with the formulations used in the analytical method hence large errors were introduced to the solution. The method is not discussed further here since not much development has evolved from it.

The mathematical basis of the PEM and UT method are further described below.

4.4.1. Point Estimate Method for PLF Studies

Point estimate methods refer to methods which require knowledge of the probability density function $f(X)$ at a specific set of values of X to get the statistical moment of $f(X)$ [4.50]. Generally, the estimation points are chosen based on the knowledge of the first few moments of the variables and eliminates finding derivatives as with analytical methods. Although PEM has been applied to other engineering fields [4.50-4.54], it was first applied to probabilistic load flow in 2005 [4.25].

Several PEM schemes (based on the number of estimation points required) have been proposed since the concept was introduced in 1975 by Rosenblueth [4.50]. This initial proposition could only handle symmetrical distributions (in fact only the normal distribution) for univariate and multivariate problems and requires 2^n estimations [4.50]. Improved methods such as the $2n$ [4.25-4.27], $2n+1$ [4.28, 4.29], $3n$ [4.28], $4n+1$ [4.28] and 5 [4.30] PEM all evolved as an

improvement in either the number of estimation points, the types of distribution to be handled or the ability to handle linear dependence between variables.

The 2*n* PEM requires *a priori* knowledge of the first three moments of all the inputs. It suffers from very large errors especially as the number of uncertainties and the variation in the input parameters increases [4.28], other PEM schemes involving more points are suggested in its place for better performance [4.28].

The 2*n*+1 PEM requires knowledge of the first four moments of all the inputs and gives better accuracy than the 2*n* scheme. Of the first four PEM schemes earlier mentioned, the 2*n*+1 scheme is the most computationally effective [4.28]. The 3*n* PEM gives exactly the same performance as the 2*n*+1 scheme when all input distributions are normally distributed, however, it results in complex estimation points when the input distribution follows a discrete distribution like the binomial distribution. Also, the 3*n* PEM technique requires a deeper knowledge of the input parameter up to the 5^{th} moment.

The 4*n*+1 requires even more moments (up to the 8^{th} moment) for the estimation points to be formed although it gives similar or slightly better performance as the 2*n*+1 PEM. One key similarity and advantage of the 2*n*+1 and 4*n*+1 is the independence of their concentration points on the number of input parameters.

Considering that the power system can consist of mixed distributions (with both continuous and discrete distribution) such as with wind power, the previously mentioned methods are not able to properly account for this as only the moments are considered in choosing the estimation points. In view of this, the 5PEM was proposed for power systems with wind penetration in [4.30]. The principles of the 2*n*+1 and the 5 PEMs are discussed below.

4.4.1.1. The 2*n*+1 PEM

As stated earlier, the 2*n*+1 PEM requires the knowledge of the first 4 moments of all input parameters. This technique has been proven to work effectively for symmetrical distributions. The estimation points referred to as “concentrations

points" and their associated significance known as the 'weights' are estimated using the first 4 moments.

For a function *Y* which is a function of *n* random variables denoted by $x_1, x_2, \ldots, x_n$ that is;

$$Y = f(x_1, x_2, \cdots, x_n) \tag{4.14}$$

Three concentration points and corresponding weights are used for each random variable with one centred on the mean, the relationship between these points and the moments are given as [4.53];

$$x_{\mathrm{i,k}} = \mu_{x_\mathrm{i}} + \xi_{\mathrm{i,k}}\sigma_{x_\mathrm{i}} \qquad i = 1,2,\cdots,n \qquad k = 1,2,3 \tag{4.15}$$

where μ_{x_i} and σ_{x_i} are the mean and standard deviation of the x_i while $\xi_{\mathrm{i,k}}$ denote the standard location given by;

$$\xi_{\mathrm{i,k}} = \frac{\lambda_{x_i 3}}{2} + (-1)^{3-k}\sqrt{\lambda_{x_i 4} - \frac{3}{4}\lambda_{x_i 3}^2} \qquad k = 1,2 \tag{4.16}$$

$$\xi_{\mathrm{i,k}} = 0 \qquad k = 3 \tag{4.17}$$

$\lambda_{x_i 3}$ and $\lambda_{x_i 4}$ are respectively the coefficient of skewness and kurtosis of the i_{th} variable.

The corresponding weights are given in (4.18) and (4.19).

$$w_{i,k} = \frac{(-1)^{3-k}}{\xi_{\mathrm{i,k}}(\xi_{\mathrm{i},1} - \xi_{\mathrm{i},2})} \qquad k = 1,2 \tag{4.18}$$

$$w_{i,k} = \frac{1}{n} - \frac{1}{\lambda_{x_i 4} - \lambda_{x_i 3}^2} \qquad k = 3 \tag{4.19}$$

For the PLF problem, each of the concentration point generated is used as the input to solve the deterministic load flow equation, while the weight of the points determines their contribution to the final moments (mean, deviation etc). With the placement of the third concentration point on the mean for all the *n* random variables, the load flow equation is evaluated $2n+1$ times instead of $3n$ times. The p_{th} raw moment of the final output parameters is given by [4.53];

$$E[y_i{}^p] = \sum_{i}^{n}\sum_{1}^{3} w_i (Y_{i,k})^p \tag{4.20}$$

where $Y_{i,k}$ is output result of the load flow studies using the k_{th} concentration point of the i_{th} variable while other variables are assumed to take their mean value. That is;

$$Y_{i,k} = f(\mu_{x_1}, \mu_{x_2}, \cdots, x_{i,k}, \mu_{x_{i+1}}, \cdots, \mu_{x_n}) \tag{4.21}$$

The central moments are easily derived from these raw moments using the relationship between them as explained in section 3.3.3 of Chapter 3.

The scheme is also able to represent correlation between the input variables by incorporating the correlation parameter into the moment of the variables using the Cholesky decomposition and Nataf transformation [4.29]. One shortfall of the scheme is its diminished accuracy for variables with mixed distributions such as the output wind power [4.30].

4.4.1.2. The 5PEM

Having identified the limitation of the 2*n*+1 scheme above, the 5PEM was specifically proposed in [4.30] because of the uniqueness of output wind power due to the variable nature of the wind speed. To fully harness the power from the wind source, the variable speed wind turbine is employed. Thus, for period with low wind speed, the wind turbine supplies no power while between the rated to cut-out wind speed, a constant wind power fixed at the rated power is produced as reflected in (4.22). With the 5PEM scheme, one concentration point is fixed each at zero output power (Region A) and the rated output power (Region C), while the other 3 points are located on the continuous part (Region B) of the wind power curve.

$$Y = \begin{cases} 0 \text{ if } X \le v_{ci} \quad \text{or} \quad X > v_{co} & \text{(Region } A\text{)} \\ \alpha + \beta X \text{ if } \quad v_{ci} \le X \le v_r & \text{(Region } B\text{)} \\ M \text{ if } \quad v_r \le X \le v_{co} & \text{(Region } C\text{)} \end{cases} \tag{4.22}$$

Y is the generated wind power while X is the wind speed at the particular instance under consideration. v_{ci}, v_{co} and v_r are respectively the cut-in, cut-out and rated wind speeds. The constants α and β are chosen such that (4.23a) and (4.23b) hold.

$$\alpha + \beta v_{ci} = 0 \qquad (4.23a)$$
$$\alpha + \beta v_{ci} = M \qquad (4.23b)$$

The weights of the concentration point located on zero and rated powers are assumed to be the probability of having zero and rated powers which is given in (4.24) and (4.25) respectively.

$$f(P_w = 0) = 1 + \exp[-\left(\frac{v_{co} - v_o}{\beta}\right)^{\alpha}] - \exp[-\left(\frac{v_{ci} - v_o}{\beta}\right)^{\alpha}] \qquad (4.24)$$

$$f(P_w = P_r) = \exp[-\left(\frac{v_r - v_o}{\beta}\right)^{\alpha}] - \exp[-\left(\frac{v_{co} - v_o}{\beta}\right)^{\alpha}] \qquad (4.25)$$

The continuous part in region B is assumed to have a probability of p_c which is given by (4.26).

$$p_c = 1 - f(P_w = 0) - f(P_w = P_r) \qquad (4.26)$$

The 3 points in the continuous part are then chosen using the process followed in the 2n+1 PEM scheme, though with a slight modification to the PDF. It is worth stating that the area under the PDF of the continuous part is not unity and a slight modification is done by dividing it by p_c to make it so. A comprehensive treatment of the 5PEM scheme is found in [4.30].

The scheme gives accurate results for the various test cases considered, even though the number of computation points increases drastically as the number of uncertain variables increases. For instance, with 10 uncertain variables within the network, the number of estimation required becomes unwieldy (close to 10 million!). Its applicability is thus restricted to problems with less than 7 uncertain variables to achieve the same number of evaluation runs with MCS (assuming 100,000 samples). The other shortfall of the scheme is the assumption of a linear relationship model for the output wind power rather than

the cubic relation (as described in section 3.5.4a). With the cubic model, a larger degree of error is noticeable while using the scheme.

4.4.2. The Unscented Transform [4.31] for PLF Studies

The unscented transform (UT) method discussed in this section follows that recently used in [4.31] for probabilistic load flow studies.

For a variable *y* related to **x** through a nonlinear function *f*,

$$y = f(\mathbf{x}) \tag{4.27}$$

The main aim is to correctly evaluate the statistical moment of the output distribution *y*. The UT method works on the fact that it is better to approximate probability distribution function rather than a nonlinear function [4.54]. With this technique, the mean ($\bar{y}$) and covariance ($\sum_{yy}$) of the output can be approximated by estimating the function at few points selected based on the mean ($\bar{\mathbf{x}}$) and covariance ($\sum_{\mathbf{xx}}$) of the input distributions. The estimation of the input variable and their corresponding weights are given in (4.28) to (4.33) [4.53]. It is worth stating that the inclusion of the covariance matrix allows for consideration of linear dependency between the variables.

$$x^0 = \bar{x} \tag{4.28}$$

$$W^0 = W^0 \tag{4.29}$$

$$x^i = \bar{x} + \left(\sqrt{\frac{\mathrm{n}}{1-\mathrm{W}^0}}\sum_{xx}\right)_{\mathrm{i}} \tag{4.30}$$

$$W^i = \frac{1-W^0}{2n} \tag{4.31}$$

$$x^{i+n} = \bar{x} - \left(\sqrt{\frac{\mathrm{n}}{1-\mathrm{W}^0}}\sum_{xx}\right)_{\mathrm{i}} \tag{4.32}$$

$$W^{i+n} = \frac{1-W^0}{2n} \tag{4.33}$$

The sum of all the weights should be unity for the above relationships to hold, also numerically efficient root finding techniques such as the Cholesky

decomposition should be used for (4.30) and (4.32) [4.31]. It is seen that the estimation points (aside from the mean) and sigma points depend on W^0, the weight of the mean point. This presents a major challenge to the technique since this control value is to be chosen arbitrarily. A value of (1/3) was suggested and employed in [4.54] since the value gave the true kurtosis for the Gaussian distribution in some cases, unfortunately this may only be applicable to problems where all the inputs are normally distributed. The main problem comes back to that of determining the appropriate value for the weight of the mean point, which if not correctly 'assumed' will lead to large error in the results [4.22].

4.5. Other Techniques

Other techniques have been proposed in the literature which combine two or more of the schemes mentioned above. The aim of most of these hybrid methods is to combine the strengths of various methods into a single technique while also trying to reduce or eliminate their drawbacks.

In [4.32], [4.33], the convolution method was combined with the Monte Carlo Simulation technique. A linearized load flow equation was employed while the MCS is used in plotting the distribution of the output functions. The Enhanced Linear Method (ELM) is employed in [4.34]. The method combines concepts from the conventional cumulant method, FFT for convolution with the point estimate also employed in calculating the mean of the outputs. The method was shown to give better results for higher moments (2^{nd} moment and above) than the $2n+1$ PEM especially when dependence exits between the random input variables [4.34].

The fuzzy load flow technique is presented in [4.35-4.38]. The technique gives the possibility distribution of the output rather than the probability distribution. The application of the method is still confined to systems with limited number of uncertainties and also yet to be applied to problems with arbitrary uncertainties like the wind power [4.57].

In [4.56], system component outages are modelled employing a probabilistic technique based on the MCS method while the load uncertainties are modelled using fuzzy sets.

The interval arithmetic load flow is employed in [4.39, 4.40] for a radial distribution system with results obtained close to that from the MCS method. Although the method performs fairly well when the interval for which the evaluation is carried out is small, but for larger intervals, the performance of the method is conservative [4.39, 4.40, 4.57].

4.6. Précis

In this chapter, probabilistic load flow techniques have been reviewed. The methods have been classified into four groups namely numerical/simulation techniques, analytical techniques, approximate techniques and hybrid schemes based on the computational principle on which they operate. The Monte Carlo Simulation technique which is a simulation scheme is the conventional means through which probabilistic load flow studies are carried out and currently still used in benchmarking the performance of new techniques due to its accuracy. However, the large sample size required before convergence is reached presents a major drawback to the method thus necessitating the proposition of other techniques.

Analytical methods such as the Convolution method, FFT technique and cumulant method were discussed as possible alternatives previously applied in solving the PLF problem. Although analytical methods are generally fast as compared with the MCS technique, they suffer from large errors due to the linearization of the load flow equation. The convolution method and the FFT technique have also been seen to involve complicated mathematics which further hampers their application.

Approximate methods briefly discussed in this chapter are viable alternatives to Monte Carlo Simulation and analytical methods. Some of the prominent approximate methods are; the point estimate method (PEM) and the unscented transform (UT) method. The PEM was further classified into 2n, 2n+1, 3n,

4*n*+1 and 5PEM based on the number of concentration points required in representing each random variable present, all the schemes require *a priori* knowledge of the moments of each random variable present. Though the 2*n* PEM required the least number of estimation points, unfortunately, it suffers from large errors as the number of uncertain variable within the system increases. The 3*n* PEM is seen to perform poorly for non-Gaussian distributions (like the binomial distribution) while the 4*n*+1 scheme is very demanding since it involves the knowledge of the first eight moments in choosing the concentration points. Of all the schemes, the 2*n*+1 PEM seemed to be the most accurate nonetheless; its higher order moments are prone to errors.

An unscented transform method whose estimation points and weights are based on the knowledge of the mean and covariance has also been examined in this chapter. Though the method presents a potential means for evaluating the first two moments of the output, its applicability is limited due to the dependence of its estimation points and weights on an arbitrary value, which then controls the accuracy of the scheme.

Few hybrid schemes previously applied to the PLF problem in the literature were also reviewed. Other uncertainty load flow techniques such as the fuzzy load flow and the interval arithmetic load flow method were also mentioned, though, their applicability is currently limited to systems with small level of uncertainties.

Considering the shortfalls of the reviewed techniques, an effective probabilistic load flow technique with the following features is desired.

i. Little computational time and burden

ii. Good level of accuracy comparable with the MCS method

iii. Applicable to practical small and large systems

iv. Able to accurately model uncertainties within the network irrespective of their distribution

v. Able to represent the dependency that exists between random uncertainties in the network.

This ideal technique could be based on providing improvement to previously discussed techniques or an entirely new method. The search for this alternative technique forms the core of this research with all the subsequent chapters in this work centred around this aim.

4.7. References

[4.1] P. Chen, Z. Chen, and B. Bak-Jensen, "Probabilistic load flow: A review," In *proc. Electric Utility Deregulation and Restructuring and Power Tech.*, pp. 1586-1591, 2008.

[4.2] P. Jogensen, J.S. Christensen and J.O. Tande, "Probabilistic Load flow calculation using Monte Carlo Techniques for Distribution Network with wind turbine," In *proc. IEEE 8th Int. Conference on Harmonics and Quality of Power*, pp. 1146-1151, 1998.

[4.3] P. Caramia, G. Carpinelli, M. Pagano & P. Varilone, "Probabilistic three-phase load flow for unbalanced electrical distribution systems with wind farms" *IET Renewable Power Generation*, vol. 1, no. 2, pp. 115–122, 2007.

[4.4] Z. Shu, and P. Jirutitijaroen, "Latin Hypercube Sampling Techniques for Power Systems Reliability Analysis with Renewable Energy Sources," *IEEE Trans. On Power Systems*, vol. 26, no. 4, pp. 2066-2073, Nov 2011.

[4.5] H. Yu, C. Y. Chung, K. P. Wong, H. W. Lee, and J. H. Zhang, "Probabilistic load flow evaluation with hybrid Latin hypercube sampling and Cholesky decomposition," *IEEE Trans. Power Systems*, vol. 24, no. 2, pp. 661–667, May 2009.

[4.6] Y. Chen, J. Wen, and S. Cheng "Probabilistic Load Flow Method Based on Nataf Transformation and Latin Hypercube Sampling" *IEEE Trans. Sustainable Energy*, vol. 4, no. 2, pp. 294-301, April 2013.

[4.7] J. F. Dopazo, O. A. Klitin, and A. M. Sasson. "Stochastic load flows," *IEEE Trans. on Power Apparatus and Systems,* PAS-94(2):2 99-303, *March* 1975.

[4.8] R. N. Allan, C. H. Griggs and M. R. G. Al-Shakarchi "Numerical techniques in probabilistic Load flow Problems," *International Journal of Numerical Methods in Engineering*, vol.10, pp. 853-860, 1976.

[4.9] R. N. Allan, A. M. Leite da Silva and R. C. Burchett "Evaluation Methods and Accuracy in Probabilistic Load Flow Solutions," *IEEE Trans. on Power Apparatus and Systems,* PAS-1000, no 5, pp. 2539-2546, May 1981.

[4.10] A. M. Leite da Silva, and R. N. Allan, "Probabilistic load flow considering network outages," *IEE proceedings*, vol. 132, no. 3, pp. 139-145, May 1985.

[4.11] A. M. Leite da Silva, and R. N. Allan, "Probabilistic load flow using multilinearisations," *IEE proceedings*, vol. 128, no. 5, pp. 280-287, Sept. 1981.

[4.12] P. S. Meliopoulos, G. J. Cokkinides, and X. Y. Chao, "A New Probabilistic Power Flow Analysis Method," *IEEE Trans. on Power Systems*, vol. 5, no. 1, pp. 182-190, Feb. 1990

[4.13] P. Zhang and S. T. Lee "Probabilistic Load Flow Computation Using the Method of Combined Cumulants and Gram-Charlier Expansion," *IEEE Trans. on Power Systems*, vol. 19, no. 1, pp. 676-682, Feb. 2004.

[4.14] A. Schellenberg, W. Rosehart, and J. Aguado, "Cumulant-Based Probabilistic Optimal Power Flow (P-OPF) With Gaussian and Gamma Distributions" *IEEE Trans. on Power Systems*, vol. 20, no. 2, pp. 773-781, May 2005.

[4.15] J. Usaola "Probabilistic load flow with wind production uncertainty using cumulants and Cornish–Fisher expansion," [Online]. Available: http://e-archivo.uc3m.es

[4.16] L. Min, and P. Zhang, "A Probabilistic Load Flow with Consideration of Network Topology Uncertainties," In *proc. The 14th International*

Conference on Intelligent System Applications to Power Systems, ISAP 2007 pp. 7-11.

[4.17] X. Wang, Y. Song, and M. Irving, *Modern Power Systems Analysis*, New York: Springer, 2008, pp. 161-178.

[4.18] M. Fan, V. Vittal, G. T. Heydt, and R. Ayyanar, "Probabilistic Power Flow Analysis With Generation Dispatch Including Photovoltaic Resources," *IEEE Trans. on Power Systems*, vol. 28, no. 2, pp. 1793-1805, May 2013.

[4.19] M. Fan, V. Vittal, G. T. Heydt, and R. Ayyanar, "Probabilistic Power Flow Studies for Transmission Systems With Photovoltaic Generation Using Cumulants," *IEEE Trans. on Power Syst*ems, vol. 27, no. 4, pp. 2251-2261, Nov. 2012.

[4.20] A. Tamtum, A. Schellenberg, and D. Rosehart, "Enhancement to the Cumulant Method for Probabilistic Optimal Power Flow Studies," *IEEE Trans. on Power Systems*, vol. 24, no. 4, pp. 1739-1746, Nov. 2009.

[4.21] O. A. Oke and D. W. P. Thomas "Enhanced Cumulant Method for Probabilistic Power Flow in Systems with Wind Generation" *In proc. 11th International Conference on Environment and Electrical Engineering (EEEIC)*, 2012, pp. 849-853.

[4.22] O. A. Oke, D. W. P. Thomas and G. M. Asher, "Enhanced Unscented Transforms method for probabilistic load flow in systems wind penetration," (submitted to IEEE Trans. on Power System)

[4.23] M. Madrigal, K. Ponnamabalam, and V.H. Quintana, "Probabilistic Optimal Power Flow," *In proc. IEEE Canadian conf. on Electrical Computer Eng.*, pp. 385-388, May 1998.

[4.24] C. Wan, Z. Xu, Z. Y. Dong, and K. P. Wong, "Probabilistic Load Flow Computation Using First-order Second-moment Method," In *Proc. Power & Energy Society General Meeting 2012*, pp. 1-6.

[4.25] C. L. Su and C. N. Lu, "Two-point estimate method for quantifying transfer capability uncertainty," *IEEE Trans. on Power Systems*, vol. 20, no. 2, pp. 573–579, May 2005.

[4.26] C. L. Su, "Probabilistic Load-Flow Computation Using Point Estimate Method," *IEEE Trans. on Power Systems*, vol. 20, no. 4, pp. 1843-1851, Nov. 2005

[4.27] G. Verbic, A. Schellenberg, W. Rosehart, and C.A. Canizares "Probabilistic Optimal Power Flow Applications to Electricity Markets," *IEEE Trans. on Power Systems*, vol. 21, no. 4, pp. 1883-1893, Nov. 2006.

[4.28] J. M. Morales and J. Pérez-Ruiz, "Point Estimate Schemes to Solve the Probabilistic Power Flow," *IEEE Trans. on Power Systems*, vol. 22, no. 4, pp. 1594-1601, Nov. 2007.

[4.29] J. M. Morales, L. Baringo, A. J. Conejo and R. Minguez, "Probabilistic Power Flow with correlated wind sources," *IET Generation Transmission Distribution*, vol.4, iss. 5, pp. 641-651, 2010.

[4.30] D.M. Outcalt. "Probabilistic Load Flow for High Wind Penetrated Power Systems based on a Five Point Estimation Method," PhD Thesis, University of Wisconsin-Milwaukee, December 2009.

[4.31] M. Aien, M. Fotuhi-Firuzabad and F. Aminifar, "Probabilistic Load Flow in Correlated Uncertain Environment Using Unscented Transformation," *IEEE Trans. on Power Systems*, vol.27, no. 4., pp 2233-2241, Nov 2012.

[4.32] A.M. Leite da Silva and V.L. Arienti "Probabilistic load flow by a multilinear simulation algorithm," *IEE Proc.*, vol. 137, no. 4, pp. 276-282, July 1990.

[4.33] A.M. Leite da Silva, V.L. Arienti, and R.N. Allan, "Probabilistic load flow considering dependence between input nodal powers," *IEEE Trans. Power Apparatus and Systems*, vol. PAS-103, no. 6, pp. 1524-1530, June 1984.

[4.34] J. Usaola, "Probabilistic load flow in systems with wind generation," [Online]. Available: http://orff.uc3m.es/bitstream/10016/8971/1/wind_plf.pdf [accessed 12/11/2012]

[4.35] V. Miranda, and J.T. Saraiva, "Fuzzy modelling of power system optimal load flow," *IEEE Trans. on Power Systems*, vol.7, no. 2, pp. 845-849, May 1992.

[4.36] M. Cortés-Carmona, R. Palma-Behnke, and G. Jiménez-Estévez, "Fuzzy Arithmetic for the DC Load Flow," *IEEE Trans. on Power Systems*, vol. 25, no. 1, pp 206-214, Feb 2010.

[4.37] M. Kim, D. Kim, Y. T. Yoon, S. Lee, and J. Park, "Determination of Available Transfer Capability Using Continuation Power Flow with Fuzzy Set Theory," In *proc. IEEE Power Eng. Society Meeting*, pp. 1-7, June 2007.

[4.38] M. A. Matos, "The Fuzzy power flow revisited," *IEEE Trans. on Power Systems*, vol. 23, no.1, pp. 213-218, Feb 2008.

[4.39] Z. Wang and F.L. Alvarado, "Interval arithmetic in power flow analysis," *In proc. Power Industry computer application conference*, pp. 156-162, May 1999.

[4.40] B. Das, "Radial distribution system power flow using interval arithmetic," *Electrical Power Energy Systems*, vol. 24, iss. 10, pp. 827-836, Dec 2002.

[4.41] A. Chaturvedi, K. Prasad, and R. Ranjan, "Use of interval arithmetic to incorporate the uncertainty of load demand for radial distribution system analysis," *IEEE Trans. on Power Delivery* vol.21, no. 2, pp. 1019-1021, April 2006.

[4.42] What is a Numerical Method?, [Online]. Available: http://www3.ul.ie/~mlc/support/CompMaths2/files/WhatNumericalMathod.pdf. [Accessed, 09/11/2012]

[4.43] R. Y. Rubinstein, *Simulation and the Monte Carlo Methods*, New Jersey: John Wiley & Sons, 1981, pp. 114-148.

[4.44] R. Green II, "Novel Computational Methods for the Reliability Evaluation of Composite Power Systems using Computational Intelligence and High Performance Computing," PhD Thesis, University of Toledo, 2012.

[4.45] R. E. Melchers, *Structural Reliability Analysis and Prediction*, Chichester: John Wiley & Sons, 2nd ed., 1999, pp. 64-93.

[4.46] W. H. Press, S. A Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in C++*, 2nd ed., New York: Cambridge University Press, 2002, pp. 308-331.

[4.47] M. D. McKay, R. J. Beckman, and W. J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 21, pp. 239–245, 1979.

[4.48] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*, 2nd ed., New York: John Wiley & Sons, 1996, pp. 104.

[4.49] Approximation theory, [Online]. Available: http://www4.ncsu.edu/~mtchu/Teaching/Lectures/MA530/chapter7.pdf

[4.50] E. Rosenblueth, ''Point estimates for probability moments.'' In p*roc., National Academy of Science*, vol. 72, iss. 10, pp. 3812–3814, 1975.

[4.51] E. Rosenblueth, ''Two-point estimates in probabilities,'' *Applied Mathematical Modelling*, vol. 5, iss.2, pp. 329–335. 1981

[4.52] K. S. Li "Point-Estimate Method for Calculating Statistical Moments," *ASCE Journal of Engineering Mechanics*, vol. 118, iss. T, pp.1506-1511, July 1992.

[4.53] H. P. Hong, "An efficient point estimate method for probabilistic analysis," *Journal of Reliability Engineering and System Safety*, vol. 59, iss. 3, pp 261–267, March 1998.

[4.54] J. T. Christian, and G. B. Baecher "Point-Estimate Method as Numerical Quadrature," *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, vol. 125, no. 9, pp. 779-786, Sept 1999.

[4.55] S. J. Julier, and J.K. Uhlmann. "Unscented Filtering and Nonlinear Estimation". *IEEE proc.,* vol. 92, no.3, pp. 401-422, March 2004

[4.56] J. T. Saraiva, V. Miranda and L. M. V. G. Pinto "Generation/transmission power system reliability evaluation by Monte Carlo Simulation assuming a fuzzy loads description," *IEEE Trans. on Power Systems,* vol. 11, iss. 2, pp. 690-695, May 1996.

[4.57] R. Raina and M. Thomas, "Fuzzy vs. Probabilistic Techniques to Address Uncertainty for Radial Distribution Load Flow Simulation," *Energy and Power Engineering*, vol. 4, pp. 99-105, 2012.

Chapter 5

The Conventional Unscented Transform Method

In this chapter, an alternative probabilistic load flow method known as the Unscented Transform method is discussed. The strength of the method in resolving some of the challenges associated with other methods reviewed in the last chapter is explored. The core mathematical theories of the method are discussed while its applicability to the probabilistic load flow problem is finally evaluated.

5.1. Mathematical Basis of the Unscented Transform Method

The Unscented Transform (UT) method was first introduced in 1997 as an alternative estimator to the Monte Carlo Simulation method [5.1]. The method has since been applied to non-linear problems in electromagnetic compatibility [5.2]-[5.6] and medical statistics [5.7] amongst other fields. In this chapter, the applicability of the UT method to the non-linear probabilistic load flow problem is explored.

In the UT method, a continuous function with probability density function (PDF) $w(x)$ is approximated as a discrete distribution w_i using deterministically chosen points referred to as sigma points (S_i) such that both distributions have the same moments. This can be mathematically represented by (5.1) while a graphical illustration is shown in Fig. 5.1.

$$E(x^k) = \int x^k w(x) dx = \sum_i w_i S_i^{\,k} \tag{5.1}$$

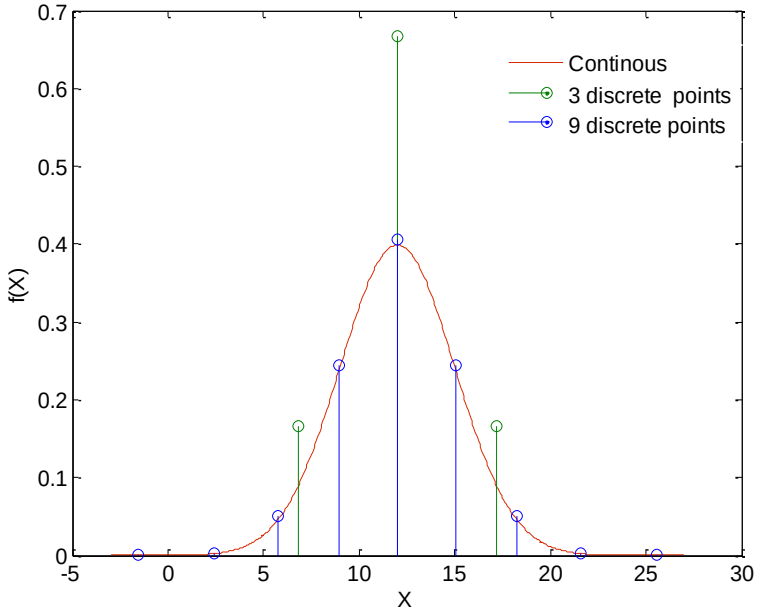

Fig 5.1: Continuous and discrete representation of a function

The initial technique for determining the sigma points and weights requires a full knowledge of the first few moments of the random variable which is similar to the approach employed in the point estimate method. Considering that (5.1) can be solved using other techniques aside from this initial proposition herewith discussed, the technique discussed in this chapter will be referred to as the conventional UT (cUT) for distinguishing purpose.

The theories of the conventional UT method for univariate and multivariate problems are expounded in the following sections. Finally, the performance of the method in solving load flow problems is evaluated using a simple 6 bus test system.

5.1.1. Conventional UT In Univariate Problems

Let x be a random variable such that it is related to another random variable y by a non-linear transformation.

$$y = f(x) \tag{5.2}$$

Let the mean of x be denoted by $\bar{x}$ and let $\hat{x}$ be a standardized variable with zero mean and the same deviation as x, that is;

$$x = \bar{x} + \hat{x} \tag{5.3}$$

Substituting (5.3) into (5.2) and applying Taylor series expansion to it gives;

$$y = f(x) = f(\bar{x}) + \hat{x}\frac{df}{d\bar{x}} + \frac{\hat{x}^2}{2!}\frac{d^2 f}{d\bar{x}^2} + \frac{\hat{x}^3}{3!}\frac{d^3 f}{d\bar{x}^3} + \cdots \tag{5.4}$$

Relating this to moments, the expectation of (5.4) $\bar{y}$ is given by;

$$\bar{y} = \mathrm{E}[f(\bar{x} + \hat{x})] = f(\bar{x}) + \mathrm{E}[\hat{x}\frac{df}{d\bar{x}} + \frac{\hat{x}^2}{2!}\frac{d^2 f}{d\bar{x}^2} + \frac{\hat{x}^3}{3!}\frac{d^3 f}{d\bar{x}^3} + \cdots] \tag{5.5}$$

For ease, let

$$g(\hat{x}) = \hat{x}\frac{df}{d\bar{x}} + \frac{\hat{x}^2}{2!}\frac{d^2 f}{d\bar{x}^2} + \frac{\hat{x}^3}{3!}\frac{d^3 f}{d\bar{x}^3} + \cdots \tag{5.6}$$

The expectation of $g(\hat{x})$ can be denoted by;

$$\mathrm{E}[g(\hat{x})] = \bar{g} \tag{5.7}$$

Substituting (5.7) into (5.5) results into a simplified representation of the function's mean as (5.8).

$$\bar{y} = f(\bar{x}) + \bar{g} \tag{5.8}$$

In like manner the variance of *y*, σ_y is given by;

$$\sigma_y^2 = \mathrm{E}[(y - \bar{y})^2] \tag{5.9}$$

$$y - \bar{y} = f(\bar{x}) + g(\hat{x}) - f(\bar{x}) - \bar{g} = g(\hat{x}) - \bar{g} \tag{5.10}$$

Substituting (5.7) and (5.10) into (5.9) gives;

$$\sigma_y^2 = \mathrm{E}[g^2(\hat{x})] - \bar{g}^2 \tag{5.11}$$

This technique can also be applied to a discrete distribution y_i in (5.12),

$$y_i = f(x_i) \tag{5.12}$$

as in (5.3), let;

$$x_i = \bar{x} + S_i \tag{5.13}$$

The mean is given as [5.5];

$$\bar{y} = \sum_{i=0}^{p} w_i f(x_i) = \sum_{i=0}^{p} w_i f(\bar{x} + S_i) \tag{5.14}$$

Since $i=0$ represents the mean point, the above can be expanded to give;

$$\bar{y} = w_0 f(\bar{x} + S_0) + \sum_{i=1}^{p} w_i f(\bar{x} + S_i) \tag{5.15}$$

From (5.1), the continuous and discrete distributions should have the same moments. Hence, (5.15) should be equal to (5.8). Following this assumption, and comparing these two equations;

$$f(\bar{x} + S_0) = f(\bar{x}) \tag{5.16}$$

From (5.16) it can be deduced that $S_0=0$.

The discrete distribution counterpart of (5.8) is given by (5.17).

$$f(\bar{x} + S_i) = f(\bar{x}) + g(S_i) \tag{5.17}$$

Substituting (5.17) into (5.15);

$$\bar{y} = w_0 f(\bar{x} + S_0) + \sum_{i=1}^{p} w_i [f(\bar{x}) + g(S_i)] \tag{5.18}$$

Simplifying (5.18) results into;

$$\bar{y} = [w_0 + \sum_{i=1}^{p} w_i] f(\bar{x}) + \sum_{i=1}^{p} w_i g(S_i) \tag{5.19}$$

To ensure the (5.19) maintains the equality between the moments of the continuous and discrete distributions, (5.20a) and (5.20b) must hold.

$$w_0 + \sum_{i=1}^{p} w_i = 1 \tag{5.20a}$$

$$\sum_{i=1}^{p} w_i g(S_i) = \bar{g} \tag{5.20b}$$

In a similar way, the variance of the discrete distribution can be evaluated.

$$y - \bar{y} = f(\bar{x}) + g(S_i) - [w_0 + \sum_{i=1}^{p} w_i] f(\bar{x}) - \sum_{i=1}^{p} w_i g(S_i) \tag{5.21}$$

Substituting (5.20a) into (5.21) gives;

$$y - \bar{y} = f(\bar{x}) - f(\bar{x}) + g(S_i) - \sum_{i=1}^{p} w_i g(S_i) = g(S_i) - \sum_{i=1}^{p} w_i g(S_i) \tag{5.22}$$

The variance is obtained by substituting (5.20b) into (5.22).

$$\sigma_y^2 = \sum_{i=1}^{\mathrm{p}} w_i [g^2(S_i) - 2\bar{g}g(S_i)] - \bar{g}^2] \tag{5.23}$$

Simplifying (5.23) gives (5.24) which is comparable to the variance for the continuous distribution (5.11).

$$\sigma_y^2 = \sum_{i=1}^{\mathrm{p}} w_i [g^2(S_i)] - \bar{g}^2 \tag{5.24}$$

To ensure (5.1) is true, (5.11) and (5.24) must be equal. Based on this, the following deductions can be made.

$$\mathrm{E}[g^2(\hat{x})] - \bar{g}^2 = \sum_{i=1}^{p} w_i [g^2(S_i)] - \bar{g}^2 \tag{5.25a}$$

$$\mathrm{E}[g^2(\hat{x})] = \sum_{i=1} w_i [g^2(S_i)] \tag{5.25b}$$

Generalizing (5.25b), the n^{th} moment of the standardized variable is given by;

$$\mathrm{E}[g^n(\hat{x})] = \sum_{i=1}^{\mathrm{p}} w_i [g^n(S_i)] \tag{5.26}$$

In general, the relationship between the sigma point and weights for the UT method and the moments of the random variable is given by [5.5];

$$\mathrm{E}[(\hat{x})^n] = \sum_{i=1}^{\mathrm{p}} w_i [(S_i)^n] \tag{5.27}$$

The moment relationship in (5.27) is employed in determining the sigma points and weights in the conventional UT method. From (5.27), it can be inferred that the accuracy of the sigma points and weights in correctly reproducing the

moments of the output distribution is dependent on the highest value of the moment order, *n*, used in choosing them. Therefore, higher moments should be included while estimating the sigma points and weights from the input distribution. Equally, the more the number of sigma points used, the easier the convergence of the discrete distribution to the continuous one.

Using (5.27), the first four moment relation is given as (5.28) for a *pth* order approximation which requires *p+1* sigma points.

$$\left.\begin{aligned} \mathrm{E}[(\hat{x})] &= \sum_{i=1}^{p} w_i[(S_i)] = 0 \\ \mathrm{E}[(\hat{x})^2] &= \sum_{i=1}^{p} w_i[(S_i)^2] = \sigma_x^2 \\ \mathrm{E}[(\hat{x})^3] &= \sum_{i=1}^{p} w_i[(S_i)^3] = \gamma_1\sigma_x^3 \\ \mathrm{E}[(\hat{x})^4] &= \sum_{i=1}^{p} w_i[(S_i)^4] = (\gamma_2 + 3)\sigma_x^4 \end{aligned}\right\} \qquad (5.28)$$

where γ_1 and γ_2 are respectively the skewness and kurtosis of the variable *x*.

Once the moments of the input variable are known, (5.28) can easily be solved using mathematical solvers like MAPLE (as employed in this work). For instance, the above was solved in [5.4] for an input following the Gaussian distribution. One possible solution for the sigma points and weights for a 4th order approximation requiring 5 points are;

$$\mathrm{S}_{1,2} = \pm\sqrt{\zeta_1}\sigma \qquad w_{1,2} = \frac{3}{4\zeta_1}$$

$$\mathrm{S}_{3,4} = \pm\sqrt{\zeta_2}\sigma \qquad w_{3,4} = \frac{3}{4\zeta_2}$$

$$S_0 = 0 \qquad w_0 = 1 - \sum_{i=1}^{4} w_i$$

$$\text{where } \zeta_{1,2} = 5 \mp \sqrt{10}$$

The final sigma points and weights follow from (5.17) as an addition of the mean value of the input random variable to the sigma points obtained above.

5.1.2. Conventional UT In Multivariate Problems

Two techniques for extending the conventional UT method to multivariate problems are discussed below. The first follows directly from the technique employed for the univariate problem above while the second technique referred to as the General Set [5.4] uniquely determines the sigma points at strategic locations while placing the variables in an Euclidean space.

5.1.2.1 Multivariate Conventional UT Using Multivariate Taylor Series Expansion

For problems with two or more random variables ($x_1,x_2,\ldots,x_n$), the general Taylor series expansion is given as (5.29).

$$y = f(x_1, x_2, \cdots, x_n) = f(\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n) + (\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_n)\frac{df(\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n)}{dx} + \cdots$$
$$\cdots \frac{1}{2!}\left[\frac{\hat{x}_1 d}{dx_1} + \frac{\hat{x}_2 d}{dx_2} + \cdots + \frac{\hat{x}_n d}{dx_n}\right]^2 f(\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n) + \cdots \qquad (5.29)$$

The relationship between the moments and sigma points are obtained in a way similar to those for the one random variable problem. The effect of the increase in the number of random variables is reflected by the increase in the number of sigma points required. The general equation representing the moment and sigma point relationship is given as [5.5];

$$\mathrm{E}[(\hat{x}_1)^a \cdots g(\hat{x}_{nrv})^k] = \sum_{i=1}^{r} w_i[({}^1S_i^a)\cdots({}^{nrv}S_i^k)] \qquad (5.30)$$

where *nrv* denotes the number of random variables while the terms ($a,\ldots,k$) define the possible combination of the random variable moments in the Taylor series expansion such that ($a+\ldots+k$) gives the order of truncation/approximation and r is one less than the total number of sigma points used.

Considering that the order of truncation of the series affects the accuracy of the output distribution, higher order terms up to the 4^{th} moment should be included

while determining the sigma points and weights to ensure accuracy of the output. To illustrate the technique, the possible moment combination for a problem involving 2 random variables will require 14 equations for a 4th order approximation of the Taylor series. This has been solved in [5.4] for a problem where all the variables follow the Gaussian distribution.

Using MAPLE, several solutions can be obtained in solving the stated problem. One possible solution for the sigma points and weights is [5.4];

$$
\begin{aligned}
&w_1 = w_2 = w_3 = w_4 = w_5 = \frac{1}{10} \\
&{}^1S_1 = 2\cos\left(0\frac{2\pi}{5}\right)\sigma_1 \qquad {}^2S_1 = 2\sin\left(0\frac{2\pi}{5}\right)\sigma_2 \\
&{}^1S_2 = 2\cos\left(1\frac{2\pi}{5}\right)\sigma_1 \qquad {}^2S_2 = 2\sin\left(1\frac{2\pi}{5}\right)\sigma_2 \\
&{}^1S_3 = 2\cos\left(2\frac{2\pi}{5}\right)\sigma_1 \qquad {}^2S_3 = 2\sin\left(2\frac{2\pi}{5}\right)\sigma_2 \\
&{}^1S_4 = 2\cos\left(3\frac{2\pi}{5}\right)\sigma_1 \qquad {}^2S_4 = 2\sin\left(3\frac{2\pi}{5}\right)\sigma_2 \\
&{}^1S_5 = 2\cos\left(4\frac{2\pi}{5}\right)\sigma_1 \qquad {}^2S_5 = 2\sin\left(4\frac{2\pi}{5}\right)\sigma_2
\end{aligned}
\tag{5.31}
$$

For problems with more random variables *nrv*, the number of required equations *NE* to be solved simultaneously depends on the order of truncation. The formula for determining *NE* for *nrv* random variables for 2nd to 6th order truncation is given in (5.32) and can be extended as the number of variables increases.

$$
NE = \begin{cases}
2^{nrv}C_1 + {}^{nrv}C_2 & \text{(2nd order)} \\
3^{nrv}C_1 + 3^{nrv}C_2 + {}^{nrv}C_3 & \text{(3rd order)} \\
4^{nrv}C_1 + 6^{nrv}C_2 + 4^{nrv}C_3 + {}^{nrv}C_4 & \text{(4th order)} \\
5^{nrv}C_1 + 10^{nrv}C_2 + 10^{nrv}C_3 + 5^{nrv}C_4 + {}^{nrv}C_5 & \text{(5th order)} \\
6^{nrv}C_1 + 15^{nrv}C_2 + 20^{nrv}C_3 + 15^{nrv}C_4 + 6^{nrv}C_5 + {}^{nrv}C_6 & \text{(6th order)}
\end{cases}
\tag{5.32}
$$

where ${}^aC_b$ denotes *a* combination *b* which is;

$${}^{a}C_{b} = \frac{a!}{(a-b)!b!}$$

Note: when $a>b$, NE terminates at the point $a=b$.

The minimum number m of sigma points needed in representing each random variable, excluding the mean point is given as [5.5];

$$m = \left\lceil \frac{NE}{nrv+1} \right\rceil \tag{5.33}$$

where $\lceil \bullet \rceil$ denotes that m takes an integer value equal or greater than the evaluated result.

As an illustration, the number of equations and minimum sigma points (Ns) for a 4th order Taylor series expansion for various number of variables is given in Table 5.1.

Table 5.1: Number of Equations and Sigma Points for Various Number of Variables

nrv	*NE*	*Ns=(m+1)*
2	14	5+1
3	34	9+1
4	69	14+1
5	125	21+1
6	209	30+1
7	329	42+1
8	494	55+1
9	714	72+1
10	1000	91+1
20	10625	506+1
100	4598125	45526+1

5.1.2.2. General Set for Multivariate Problems [5.4]

In [5.4], a simple technique was proposed for problems with non-uniform random variables. The scheme defines the random variables in an Euclidean space such that the sigma points are located on the edges and axis of an *nrv*-dimensional cube. With the general set, the number of sigma points is evaluated using $2^{nrv}+2nrv$. The sigma points (2^{nrv}) located on the edges is given by;

$$S_e = (\pm 1, \cdots, \pm 1)\frac{\sqrt{nrv+2}}{\sqrt{nrv}} \tag{5.34a}$$

$$W_e = \frac{nrv^2}{2^{nrv}(nrv+2)^2} \tag{5.34b}$$

While those located on the axis S_a and W_a are given by;

$$S_a = (\pm 1, \cdots, 0)\sqrt{nrv+2} \tag{5.35a}$$

$$W_a = \frac{1}{(nrv+2)^2} \tag{5.35b}$$

With the general set, the total required sigma points and weights are more than those presented in Table 5.1. A picture of the sigma point growth is presented in [5.4].

5.2. Case Studies and Discussion

The aim of this section is to demonstrate the applicability of the conventional UT method in solving probabilistic load flow problems. This will be achieved using 4 different case scenarios. For all the cases, a simple 6 bus IEEE test system (Appendix C) with random variable(s) introduced on some buses is used. In line with previous research, the performance of the conventional UT method is appraised by comparing results obtained with those from the Monte Carlo Simulation (MCS) method. In addition, the conventional UT method is also compared against existing methods like the point estimate method (PEM) and the cumulant method (CM).

The percentage error index [5.9] in the moments of the output results with reference to the MCS is used in assessing the performance of the conventional UT method.

$$\varepsilon_{\mu}^{u} = \left| \frac{\mu_{MCS}^{u} - \mu_{UT/PEM/CM}^{u}}{\mu_{MCS}^{u}} \right| \times 100\% \tag{5.36}$$

where u represents the variable (e.g. voltage, active power etc.) and μ stands for the moments.

In each of the cases, 100,000 samples were generated for the MCS method as with this, we are 95% confident that the average percentage errors in the active power flow is less than 0.3% for cases where all random variables are normally distributed while it is less than 6.5% for the non-normal variable cases (Table 5.2). It is worth stating that the accuracy of the MCS method increases as the number of samples increases. The program for all the methods was implemented using MATLAB while MATPOWER [5.10] was incorporated to run the load flow.

Table 5.2: Average Error in the Mean Active Power Flow Using 100,000 Samples

Case Number	Average Error using 100,000 Samples for MCS (%)
Case 5.1	0.09143
Case 5.2	5.69658
Case 5.3	6.49687
Case 5.4	0.23712

5.2.1. Case 5.1: Univariate Normally Distributed Random Variable

The focus of the study is to evaluate the performance of the cUT method while treating problems with univariate symmetrical random variables. The study also illustrates the level of accuracy achievable with the cUT method as the number of sigma point increases.

The standard 6 bus test system was modified such that the active mean power on bus 3 is assumed to be normally distributed with a coefficient of variation of

10%. Using (5.36), the average percentage errors up to the 4th moment for the cUT method with 2, 3 and 5 sigma points are presented in Table 5.3. From the results, it is clear that the 2 sigma point cUT is only accurate up to the second moment while the 3 and 5 sigma points cUT gave errors less than 3% for the first two moments. However, the 5 sigma point cUT consistently gives accurate results as compared with the 3 sigma points cUT. The CDF of the voltage on bus 4 is shown in Fig. 5.2. From the inset, it is seen that the 5 sigma point cUT gave the best fit to the MCS method. The plot from the 2 point cUT performance was the poorest due to its inability to accurately estimate the higher order moments required to plot the curve using the Cornish Fisher series as discussed in section 3.4. of Chapter 3. In view of this, it is recommended that a minimum of 3 sigma points should be used in discretizing the continuous distribution to have a fairly accurate result for the distribution plot.

Table 5.3: Average Percentage Error Indices for Case 5.1

Moments		2 Pt. cUT	3 Pt. cUT	5 Pt. cUT	PEM	CM
Mean	V	1.97e-4	2.07e-4	2.07e-4	2.07e-4	6.63e-3
	δ	0.00404	0.00412	0.00412	0.00412	0.03770
	P*i,j*	0.00603	0.00607	0.00607	0.00607	0.01654
	Q*i,j*	0.00704	0.00724	0.00724	0.00724	0.16925
Std	V	0.11471	0.31030	0.31119	0.31030	0.10500
	δ	0.25161	0.31316	0.31241	0.31316	0.22410
	P*i,j*	0.29212	0.31445	0.31454	0.31445	0.09224
	Q*i,j*	2.20980	0.30912	0.30912	0.30912	2.47863
Skewness	V	100.000	2.96522	2.43129	2.96522	100.000
	δ	100.000	7.68569	7.23103	7.68569	100.000
	P*i,j*	100.000	34.9859	34.6633	34.9859	100.000
	Q*i,j*	100.000	5.62707	5.11585	5.62707	100.000
Kurtosis	V	67.2905	1.65362	0.21310	1.65362	1.98491
	δ	66.6997	0.07337	0.25040	0.07337	0.09916
	P*i,j*	66.5865	0.24381	0.34064	0.24381	0.24048
	Q*i,j*	70.4030	9.36145	0.51828	9.36145	11.3181

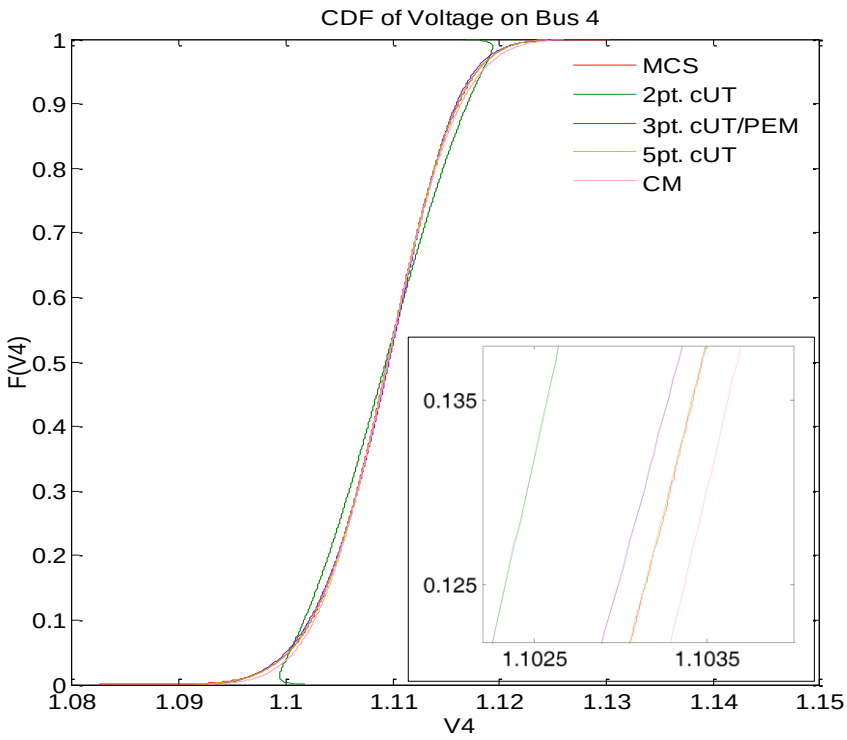

Fig 5.2: CDF plot of Voltage on Bus 4 (Detail shown in inset)

Equally, to rightly place the performance of the cUT, the results are also compared against those from the PEM and CM as shown in Table 5.3 and Fig. 5.2. From this result, it is seen that the PEM gives exactly the same result as the 3 sigma point cUT. This is as expected since both methods require 3 estimation points which have been both derived using the moments of the input variable. From the overall results, it is seen that the cUT method gives better results as compared with the CM. Considering the computation time for each method in Table 5.4, it is seen that CM requires the least amount of time of all the methods since only one computation is carried out.

Table 5.4: Computation Time For Methods in Case 5.1

Method	MCS	2 Pt. cUT	3 Pt. cUT	5 Pt. cUT	PEM	CM
Computation Time (sec)	750.912	0.055627	0.061585	0.074164	0.061585	0.028434

5.2.2. Case 5.2: Univariate Non-Gaussian Random Variable

In this study, the performance of the cUT method for problems involving single non-normal random variable is evaluated. Case 5.1 is adjusted such that the load on Bus 3 is replaced with a 55MW rated wind farm. The wind speed is assumed to follow the Weibull distribution whose parameters [5.11] are given in Table 5.5. The wind turbine parameters are also shown in the table. The cumulative wind power is fed into the system as a negative load.

Table 5.5: Wind Turbine and Wind Speed Parameters

Rated Power (MW)	Cut-in speed (m/s)	Rated wind speed (m/s)	Cut-out speed (m/s)	Wind Shape parameter	Wind Scale parameter
1	4	15	25	3.97	10.7

The average percentage error indices obtained as compared with the MCS (100,000 runs) are presented in Table 5.6. It is worth stating that the 2 points cUT method was not used considering that the wind power model as discussed in section 3.5.4 (Chapter 3) requires partitioning the wind power output for variable wind speed turbines. As such, a sigma point is fixed at zero and another at the rated output wind power, hence the least number of sigma points that can be used is 3. In view of this, the 3 points cUT uses one sigma point (the mean) from the continuous part while 3 points are chosen from the continuous part for the 5 points cUT. For the PEM, the method in [5.12] which considers the wind speed rather than the wind power as the random variable is used.

From the results in Table 5.6, the inefficiency of the 3 points cUT is noticeable. This is due to the earlier stated fact regarding the sigma point selection. The 5 point cUT gives the best result overall. The cumulant method which involves a single computation also gives fairly accurate result. The error in the cumulant method can be attributed to the output wind distribution having active points away from the mean value. To resolve this, an enhancement to the cumulant

method has been proposed in [5.13]. Generally, it is seen that the results presented in Table 5.3 for the normal random variable case are better than those in Table 5.6 due to the non-symmetrical nature of the output wind power distribution. The accuracy of the methods is thus affected by the type of (random variable) distribution under consideration.

Table 5. 6: Average Percentage Error Indices for Case 5.2

Moments		3 Pt. cUT	5Pt. cUT	PEM	CM
Mean	V	0.02057	8.08054e-5	0.06756	0.02580
	δ	0.31494	0.03421	1.05818	0.37482
	P*i,j*	2.97800	0.27137	8.27511	3.85480
	Q*i,j*	0.78000	0.00956	0.25623	0.97783
Std	V	61.1553	0.11429	10.9163	14.0502
	δ	54.9916	0.08715	5.78162	0.41455
	P*i,j*	54.6800	0.08626	5.98060	0.54953
	Q*i,j*	53.5722	0.09352	7.06495	19.6638
Skewness	V	673.827	33.4753	232.885	191.312
	δ	391.603	0.36885	18.0957	1.89796
	P*i,j*	386.971	0.28718	18.6556	2.73878
	Q*i,j*	459.132	27.9033	71.7575	139.917
Kurtosis	V	1338.81	9.13140	38.9689	48.1847
	δ	893.904	0.21952	6.26801	0.96885
	P*i,j*	886.411	0.66233	6.89060	3.09220
	Q*i,j*	968.017	8.48978	30.4196	37.7007

To further reflect the performance of the methods, the CDF plot of the voltage on Bus 4 is shown in Fig. 5.3. The plot using the 3 point cUT method has been intentionally left out due to the huge error in its higher order moment which resulted in a wide distortion of the plot. The plot (as seen in the inset) affirms the accuracy of the 5 point cUT since it gave the closest fit to the MCS method. The large variation in the PEM is as a result of its poor performance at evaluating the higher order moment for non-symmetrical distributions like that of the wind power.

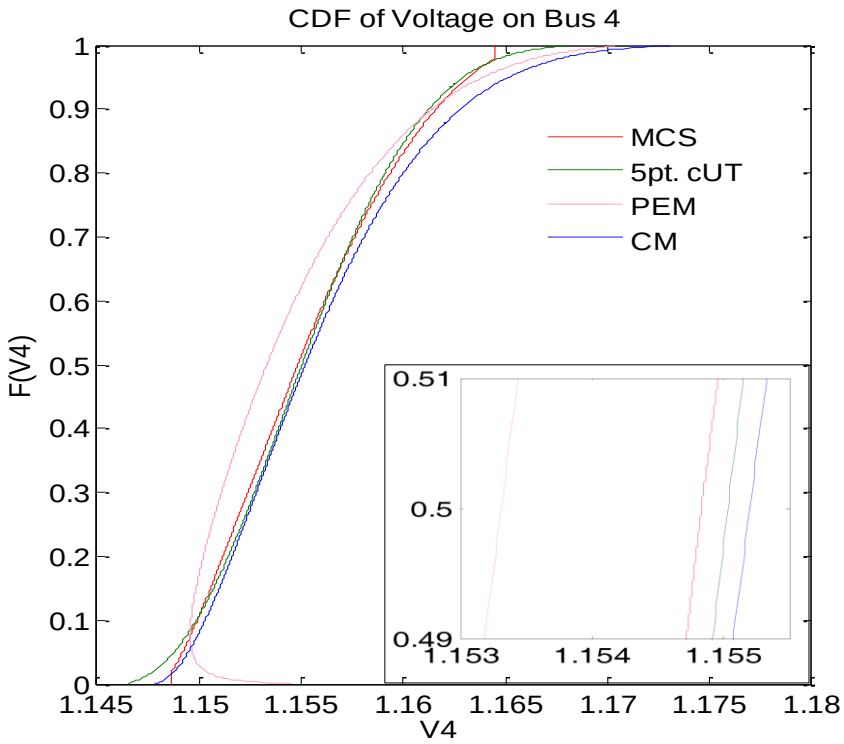

Fig 5.3: CDF Plot for Voltage on Bus 4 for Case 5.2 (Detail Shown in inset)

5.2.3. Case 5.3: Multivariate Normally Distributed Variable Problem

In this section, the ability of the cUT techniques discussed in sections 5.1.2.1 and 5.1.2.2 are evaluated. To test this, the previous case is modified; the active power on buses 3 and 5 have been assumed to be normally distributed with both having a 10% coefficient of variation.

The average percentage error indices are presented in Table 5.7. From the results, it is seen that the General Set cUT gives slightly better results than the Taylor series based multivariate cUT (denoted T.S.M. in the table), nonetheless, the latter requires more computational time as presented in Table 5.8. The large percentage error in the skewness using the cumulant method is noticeable. This is due to the skewness and kurtosis of the normal distribution which are used in the CM method being zero, thus, producing 100% percentage error relative to the MCS value. Results obtained for the PEM follow the same trend as those in [5.12]. Though about the same number of computation is required for both the PEM and T.S.M cUT, the T.S.M cUT techniques performed better than the PEM.

Table 5.7: Average Percentage Error Indices for Case 5.3

Moments		T.S.M	General Set	PEM	CM
Mean	V	0.00104	0.00104	0.00105	0.01070
	δ	0.02679	0.02679	0.02687	0.10915
	P*i,j*	0.03420	0.03419	0.03421	0.04457
	Q*i,j*	0.03913	0.03912	0.03925	0.24871
Std	V	0.06968	0.06968	0.12122	0.27158
	δ	0.02936	0.02950	0.03705	0.09800
	P*i,j*	0.06712	0.06704	0.06733	0.17777
	Q*i,j*	0.04935	0.05255	0.39507	1.44371
Skewness	V	2.25649	1.53661	43.4231	100.000
	δ	5.23306	4.13746	58.1719	100.000
	P*i,j*	10.2480	9.33190	81.7092	100.000
	Q*i,j*	6.64280	3.42245	80.2789	100.000
Kurtosis	V	2.70175	1.68679	30.2088	2.04714
	δ	1.11309	0.58099	36.4028	0.64248
	P*i,j*	0.83869	0.58995	30.2283	0.58742
	Q*i,j*	6.94073	5.27384	22.3406	5.45092

Table 5.8: Computation Time for Methods in Case 3

Method	MCS	T.S.M	General Set	PEM	CM
Number of Evaluations	1E5	6	9	5	1

In the scenario presented, it is seen that both cUT techniques efficiently handled the 2 random variable case. From Table 5.1, considering the number of equations to be solved simultaneously, the Taylor series based multivariate cUT becomes computationally inefficient and practically unrealistic especially in the modern day power system where the number of uncertainties is high. To illustrate this, a problem involving 100 random variables (typical of power

systems) will require solving close to 5 million equations simultaneously with each random variable represented using over 45000 sigma points. Similarly, the general set will require over 1.26×10^{30} for the same problem.

5.2.4. Case 5.4: Multivariate Problem with Normal and Non-Normal Variables

This section further evaluates the performance of the multivariate conventional UT techniques. As with the previous case study, 2 random variables are included in the 6 bus test system. However, the active load on bus 3 was replaced with the wind farm output used in case 5.2 (Section 5.2.2) while the active load demand on bus 5 is assumed to follow the normal distribution with a 10% coefficient of variation.

Table 5.9: Sigma Points Using the General Set

Sigma Point for Bus 3 (MW)	Sigma Point for Bus 5 (MW)
36.5202	34.2426
-1.63029	34.2426
36.5202	25.7574
-1.63029	25.7574
17.4449	36.0000
44.4214	30.0000
17.4449	24.0000
-9.53152	30.0000
19.0752	30.0000

The sigma points obtained using the general set are presented in Table 5.9. From the results, it is seen that the general set performs poorly in determining the sigma points for the output wind power due to the negative values. This is because the sigma points for the general set are determined using the mean and standard deviation of the variable only. This makes the general set inefficient for non-symmetrical distributions with very active higher order moments.

Because of the negative values as shown in Table 5.9, the general set was not employed in carrying out the probabilistic load flow.

For the Taylor series based multivariate cUT, equation (5.29) was employed using the moments of the variables. Disappointingly, the 14 equations needed to be solved could not be resolved within 4312.7 seconds using MAPLE. This time is more than that required by the MCS!

From this, it is clear that both conventional UT method discussed in this chapter are unable to correctly treat non-normal multivariate problems, therefore an alternative approach to the conventional UT method is desired.

5.3. Précis

In this chapter the conventional Unscented Transform has been presented as a technique for solving the probabilistic load flow problem. The accuracy of the method in dealing with both normally and non-normally distributed univariate and multivariate distribution load flow problems have been explored. The results obtained have been compared with those from other 3 methods namely the point estimate method, the cumulant method and the Monte Carlo Simulation method which is the benchmark.

The conventional Unscented Transform method is seen to perform accurately in treating univariate normal random variable problems while its efficiency diminishes for problems involving non-symmetrical random variable such as the output wind power for the variable speed turbine.

Two techniques were also presented in treating multivariate problem. The first is based on Taylor series expansion for multivariable while the other referred to as the General Set is based on a geometric approach which represents the random variables in an Euclidean space. For the multivariate normal distributed case, a fairly good level of accuracy was obtained using both techniques with the general set performing slightly better. Unfortunately, a prohibitive number of sigma points is required using both techniques as the number of uncertain random variables increase. This makes them impractical for power systems with large number of uncertainties. Furthermore, the two

techniques were unsuccessful in solving multivariate problems with non-symmetrical distributions like the wind power distribution.

In view of the identified problems with the conventional Unscented Transform Technique, an alternative method capable of efficiently managing both symmetrical and asymmetrical distributed random variable is desired. One possible solution to this is to solve (5.1) using alternative approaches or completely using a different technique. The former is discussed in the next chapter.

5.4. References

[5.1] S. J. Julier, and J. Uhlmann. "Consistent Debiased Method for Converting Between Polar and Cartesian Coordinate Systems," *proc. of the 1997 SPIE Conference on Acquisition, Tracking, and Pointing*, vol. 3086, pp. 110-121.

[5.2] J. B. J. Pereira, L. de Menezes, and G. A. Borges, "Statistical Analysis of Induced Ground Voltage using the TLM+UT Method," *IEEE International symposium on Electromagnetic Compatibility, Hamburg*, 8-12 Sept. 2008.

[5.3] L. R. A. X. de Menezes, D. W. P. Thomas, C. Christopoulos, A. Ajayi and P. Sewell, "The Use of Unscented Transforms for Statistical Analysis in EMC," *IEEE International symposium on Electromagnetic Compatibility, Hamburg*, 8-12 Sept. 2008.

[5.4] L. R. A. X. de Menezes, A. Ajayi, C. Christopoulos, P. Sewell and G. A. Borges, "Efficient Computation of Stochastic Electromagnetic Problems Using Unscented Transforms," *IET Science, Measurement and Technology*, vol. 2, iss. 2, pp. 88-95, March 2008

[5.5] A. Ajayi, "Direct Computation of Statistical Variations in Electromagnetic Problems," PhD Thesis, University of Nottingham, May 2008.

[5.6] D. W. P. Thomas, O. A. Oke, L. R. A. X. de Menezes and C. Christopoulos, "The Use of Unscented Transforms in Modelling the Statistical Response of Nonlinear Scatterer in a Reverberation

Chamber," in *proc. URSI GASS 2011, Istanbul Turkey, 13-20 August 2011*

[5.7] I. dos Santos, D. Haemmerich, D. Schutt, A. F. da Rocha and L. R Menezes, "Probabilistic finite element analysis of radio frequency liver ablation using unscented transform," *Physics in Medicine and Biology*, vol. 54, No. 3, pp. 627-640, Jan. 2009

[5.8] J. Usaola, "Probabilistic Load Flow in Systems with high wind power penetration," [Online]. Available: *http://e-archivo.uc3m.es/bitstream/10016/2845/1/stochLF.pdf*

[5.9] J. M. Morales and J. Perez-Ruiz, "Point Estimate Scheme to solve the Probabilistic Power Flow," *IEEE Trans. Power Systems*, vol. 22, No 4, pp. 1594-1601, Nov. 2007.

[5.10] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER's Extensible Optimal Power Flow Architecture," *Power and Energy Society General Meeting, 2009 IEEE*, pp. 1-7, July 26-30 2009.

[5.11] Z. Bie, G. Li, H. Liu, X. Wang and X. Wang, "Studies on Voltage Fluctuation in the Integration of Wind Power Plants Using Probabilistic Load Flow," *presented at the Power and Energy Society General meeting, Pittsburgh*, pp. 1-7, 20-24 July 2008.

[5.12] J. M. Morales, L. Baringo, A. J. Conejo and R. Minguez, " Probabilistic Power Flow with Correlated Wind Sources," *IET Generation Transmission and Distribution*, vol. 4, iss. 5, pp. 641-651, 2010

[5.13] A. Tamtum, A. Schellenberg, and D. Rosehart, "Enhancement to the Cumulant Method for Probabilistic Optimal Power Flow Studies," *IEEE Trans. Power Systems*, vol. 24, no 4, pp. 1739-1746, Nov. 2009.

Chapter 6

Gaussian Quadrature for Unscented Transform Method

In this chapter, an alternative approach of viewing the unscented transform equation is presented. The method treats the unscented transform (UT) equation as an integration problem using the Gaussian quadrature technique. The mathematical basics of the Gaussian quadrature principle are first expounded while the practicability of the method is tested in the latter part.

6.1. Gaussian Quadrature: Basics

Gaussian quadrature is an approximate technique which estimates the definite integral of a function over a given interval. The fundamental theorem of Gaussian quadrature states that "the optimal abscissas of an *N*-point Gaussian quadrature formula are precisely the root of the orthogonal polynomial for the same interval and weighting function" [6.1, 6.2]. Gaussian quadrature technique is optimal because it exactly fits the polynomial up to *2N-1* degree [6.1]. This theorem can be applied to the UT equation in (5.1) which is repeated as (6.1)

$$E(x^k) = \int x^k w(x)dx = \sum_i w_i S_i^{\,k} \tag{6.1}$$

Let $P_l(x)$ be an l degree nontrivial polynomial orthogonal to the weighting function $w(x)$. Following the principle of orthogonality, then, the integral of (6.1) over the interval a and b becomes;

$$\int_a^b x^k w(x) P_l(x)dx = 0 \tag{6.2}$$

From the theorem, the zeros of the polynomial $P_l(x)$ produce l-abscissas x_i (which corresponds to the sigma points S_i). The corresponding l-weights w_i are obtained as [6.2, 6.8];

$$w_i = \frac{1}{P_l'(x_i)} \int_a^b \frac{P_l(x)}{x - x_i} dx \tag{6.3}$$

where P_l' is the first differential of P_l.

For weighting functions (probability density functions) whose orthogonal polynomials are known, the above technique can be easily applied. For instance, a Gaussian distributed weighting function is orthogonal to the Hermite polynomial while the exponential distribution is orthogonal to the Laguerre polynomial. Examples of other classical orthogonal polynomials and their corresponding weighting functions are found in [6.3, 6.4]. For arbitrary distributions with no known classical orthogonal polynomial, the corresponding orthogonal polynomial has to be generated in order to apply this principle.

6.2. Orthogonal Polynomial Generation

Two functions *f(x)* and *g(x)* are said to be orthogonal in an interval [*a, b*] over a weighting function *w(x)* if their inner product is zero, while they are orthonormal if the inner product is unity. The inner product relationship is presented in (6.4).

$$\langle f, g \rangle = \int_a^b w(x) f(x) g(x) dx = \begin{matrix} 0 & if & f(x) \neq g(x) \\ 1 & if & f(x) = g(x) \end{matrix} \tag{6.4}$$

All orthogonal polynomial satisfy a 3-term recurrence relation which is premised on the shift property in the inner product [6.4] i.e.

$$(xf, g)_{dW} = (f, xg)_{dW} \qquad \text{for all } f, g \in \mathbf{P} \tag{6.5}$$

where $dW = w(x)dx$ is the induced positive measure.

This recurrence relation allows for easy generation of the orthogonal polynomial and also the easy computation of the polynomial's root as the eigenvalues of a symmetrical tridiagonal Jacobian matrix [6.4] amongst other things.

The recurrence relation is mathematically represented as [6.1];

$$\begin{aligned} &P_{l+1}(x) = (x - a_l)P_l(x) - b_l P_{l-1}(x) \qquad l = 0,1,2,... \\ &P_{-1}(x) = 0 \\ &P_0(x) = 1 \end{aligned} \tag{6.6}$$

where a_l and b_l are the recurrence coefficients given by;

$$a_l = \frac{\langle xP_l, P_l \rangle dW}{\langle P_l, P_l \rangle dW} \qquad l = 0,1,2,... \tag{6.7}$$

$$\begin{aligned} b_l &= \frac{\langle P_l, P_l \rangle dW}{\langle P_{l-1}, P_{l-1} \rangle dW} \qquad l = 1,2,... \\ b_0 &= \langle P_0, P_0 \rangle dW \qquad l = 0 \end{aligned} \tag{6.8}$$

and $P_l(x) = P_l(x; dW)$, $l = 0,1,2,...$ is a set of monic polynomials with respect to the measure *dW*.

For this case where the weighting function is non-classical and the recurrence coefficient unknown, one way to build the recurrence relation involves the prior knowledge of the first *2l* moments of the weighting function [6.1]. Using the relationship in (6.9), a_l and b_l can be obtained for building the recurrence relation in (6.6).

$$\mu_l = \int_a^b x^l W(x) dx \qquad l = 0,1,\cdots,2N-1 \tag{6.9}$$

Unfortunately, the results produced using this procedure is extremely ill-conditioned even for functions with classical weights, as such the procedure is not applicable [6.1, 6.4] for problems with arbitrary weighting functions. In view of this limitation, a robust approach using a discretization scheme will be applied.

In discretization schemes, the weighting function is approximated by discrete points which are subsequently used for the computation of the recurrence coefficients. The discretization techniques which employ discrete measure and discrete points work perfectly once the discrete measure converges to the continuous one. The Stieltjes procedure and the Lanczos-type algorithm are

examples of such discretization schemes [6.4]. The Stieltjes procedure is further discussed since it requires less execution time as compared with the Lanczos type algorithm [6.4].

6.2.1. Stieltjes Procedure

The Stieltjes procedure provides a way of computing the recurrence relation of a discrete measure. The procedure involves representing the weighting function dW by an M-point discrete measure dW_M using a suitable interpolatory quadrature rule.

The interpolatory quadrature helps in approximating the definite integral of a given function f by a weighted sum such that [6.5], [6.8];

$$\int_{-1}^{1} f(x)dx = \sum_{k=-1}^{n} w_k f(x_k) \tag{6.10}$$

Considering that our aim is to solve for the recurrence relation terms (6.6) using (6.7)-(6.8), the objective at this stage is to solve for the inner product. To avoid singularities, the evaluation points are chosen away from the endpoints [6.4]. Applying (6.10) to the inner product definition in (6.4) gives;

$$\int_{-1}^{1} w(x) f(x) g(x) dx \cong \sum_{k=1}^{n} w_k w(x_k) f(x_k) g(x_k) \tag{6.11}$$

Equidistance interpolatory rules such as the Newton Cotes and its variants (Simpson rule, Trapezoidal rule and midpoint rule) all results into large errors and can diverge as fast as 2^n even for smooth functions [6.1, 6.4, 6.6]. Hence, the Fejér's first quadrature rule which is similar to the Clenshaw-Curtis rules is applied.

With Fejér's first rule, the interpolation points are chosen as the zeros of the Chebyshev polynomial of first kind such that the nodes and weights are obtained using (6.12a, 6.12b).

$$x_k = \cos\theta_k, \qquad \theta_k = \frac{(2k-1)\pi}{2n}, \qquad k = 1:n \tag{6.12a}$$

$$w_k = \frac{2}{n}\left(1 - 2\sum_{j=1}^{\lfloor n/2 \rfloor}\frac{\cos(2j\theta_k)}{4j^2-1}\right), \qquad k = 1:n \tag{6.12b}$$

where $\lfloor n/2 \rfloor$ is the largest integer less than or equal to *n/2*.

Considering that (6.11) is defined within the canonical interval [-1,1], a transformation equation (6.13) [6.7] is employed in mapping it to any arbitrary interval [a,b].

$$\int_a^b w(x) f(x) g(x) dx = \int_{-1}^{1} w(\phi(\tau)) f(\phi(\tau)) g(\phi(\tau)) \phi'(\tau) d\tau \tag{6.13}$$

The linear transformation (6.14) employed in [6.4] for finite [a,b] is used in this work.

$$\phi(\tau) = \begin{cases} \frac{1}{2}(b-a)\tau + \frac{1}{2}(b+a) & \text{if } -\infty < a < b < \infty \\ b - \frac{1-\tau}{1+\tau} & \text{if } -\infty = a < b < \infty \\ a + \frac{1+\tau}{1-\tau} & \text{if } -\infty < a < b = \infty \\ \frac{\tau}{1-\tau^2} & \text{if } -\infty = a < b = \infty \end{cases} \tag{6.14}$$

The points obtained using the Fejer interpolatory rule are subsequently used in computing the discrete inner product and hence the discrete recurrence coefficients $a_{l,M}$ and $b_{l,M}$. These recurrence coefficients are easily computed following this sequence;

$$P_{0,M} \rightarrow a_{0,M}, b_{0,M} \rightarrow P_{1,M} \rightarrow a_{1,M}, b_{1,M} \rightarrow \ldots$$

The process is continued until all the coefficients are obtained.

This method proves to be accurate since the discrete recurrence coefficients converge to the continuous ones as $m \rightarrow \infty$. The Stieltjes procedure is applicable not only to a continuous distribution but also to mixed distributions as demonstrated in [6.4]. The only modification is the addition of the discrete

component to the discretized continuous part. A detailed explanation of the Stieltjes procedure can be found in [6.4], [6.7].

6.3. Generation of Sigma Points and Weights for Univariate Problems

As previously mentioned, the recurrence relation allows for computation of the polynomial's roots as the eigenvalue of the tridiagonal Jacobi matrix given in (6.15) [6.1], [6.4]. The recurrence relation coefficient obtained using the Stieltjes procedure are substituted into (6.15) to determine the sigma points.

$$J_n = \begin{bmatrix} a_o & \sqrt{b_1} & 0 & 0 & \cdots & 0 & 0 \\ \sqrt{b_1} & a_1 & \sqrt{b_2} & 0 & \cdots & 0 & 0 \\ 0 & \sqrt{b_2} & a_2 & \sqrt{b_3} & \cdots & 0 & 0 \\ 0 & 0 & \sqrt{b_3} & \ddots & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \sqrt{b_{n-2}} & 0 \\ 0 & 0 & 0 & 0 & \sqrt{b_{n-2}} & a_{n-2} & \sqrt{b_{n-1}} \\ 0 & 0 & 0 & 0 & 0 & \sqrt{b_{n-1}} & a_{n-1} \end{bmatrix} \tag{6.15}$$

The corresponding weights are computed using

$$W_l = b_0 v_{l,1}^2 \qquad l = 1,...,n \tag{6.16}$$

where $v_{l,1}$ is the first component of the normalized eigenvector corresponding to *the lth* eigenvalue of J_n.

6.4. Sigma Points and Weights Generation in Multivariate Problems

The process described above works efficiently for single variable problems involving one dimensional integration. But for a multivariate problem with *n* uncertainties, the integration is carried out *n* times.

For a multivariate problem, (6.1) is modified as;

$$E(\mathbf{x}^k) = \int \cdots \int \mathbf{x}^k w(\mathbf{x}) d\mathbf{x} \tag{6.17}$$

where vector $\boldsymbol{x} = \{x_1, x_2, \ldots, x_n\}$.

This process is tedious since it requires a large number of integrations to be carried out. Nonetheless, for cases where the variables in **x** can be factored out

independently of each other, the integration becomes simple and can be approximated as the product of the one dimensional integrals of the variables. The above can be rearranged using tensor product [6.11] to give (6.18).

$$E(\mathbf{x}^k) = \int (x_1^k w(x_1) dx_1 \cdots \int (x_n^k w(x_n) dx_n = \cdots$$
$$\cdots \sum_{i1} \cdots \sum_{in} (w_{i1} \otimes \cdots \otimes w_{in})(S_{i1}, \cdots, S_{in}) \qquad (6.18)$$

Using the above, the sigma points for each variable are evaluated using the univariate scheme in the previous section. The final sigma points and weights are then obtained as the product of the individual weights and sigma points. As an illustration, the final sigma points and weights for a multivariate problem involving two random variables with each discretized by 5 points is chosen in Table 6.1.

Table 6.1: Sigma Point and Weight Generation for a Bivariate Problem

Sigma Points / Weights	$^2S_0$	$^2S_1$	$^2S_2$	$^2S_3$	$^2S_4$
$^1S_0$	$^1W_0{}^2W_0$	$^1W_0{}^2W_1$	$^1W_0{}^2W_2$	$^1W_0{}^2W_3$	$^1W_0{}^2W_4$
$^1S_1$	$^1W_1{}^2W_0$	$^1W_1{}^2W_1$	$^1W_1{}^2W_2$	$^1W_1{}^2W_3$	$^1W_1{}^2W_4$
$^1S_2$	$^1W_2{}^2W_0$	$^1W_2{}^2W_1$	$^1W_2{}^2W_2$	$^1W_2{}^2W_3$	$^1W_2{}^2W_4$
$^1S_3$	$^1W_3{}^2W_0$	$^1W_3{}^2W_1$	$^1W_3{}^2W_2$	$^1W_3{}^2W_3$	$^1W_3{}^2W_4$
$^1S_4$	$^1W_4{}^2W_0$	$^1W_4{}^2W_1$	$^1W_4{}^2W_2$	$^1W_4{}^2W_3$	$^1W_4{}^2W_4$

The number of computations (*Nc)* required using the above is high as indicated in (6.19). This indicates an exponential growth in the number of computations to be carried out as the dimension increases.

$$Nc = N^d \qquad (6.19)$$

where *N* is the number of sigma points used in discretizing each variable and *d* the dimension of the problem (which depends on the number of uncertain parameter in the problem).

6.5. Case Studies

This section discusses the applicability of the Gaussian Quadrature based Unscented Transform to the modern day power system. Since this technique is

proposed with the aim of overcoming the problems identified with the conventional UT method in the previous chapter, its accuracy in dealing with univariate and multivariate problems involving symmetrically and non-symmetrically distributed variables will be evaluated. To achieve the above, four case studies will be considered. The first three cases have been adopted from the previous chapter while the fourth case examines the performance of the method in a slightly larger power system with more uncertainties.

To adequately evaluate the performance of the technique, results have been compared with those from the Monte Carlo Simulation (MCS) method, the conventional UT (cUT), the Cumulant method (CM) and the point estimate method (PEM). For clarity, the Gaussian Quadrature based UT will be referred to as gUT. As in the previous chapter, 100,000 samples are used for the MCS method. To properly assess the performance of the gUT method, the average percentage indices relative to the MCS method is estimated using (5.36). The programme for all methods have been implemented using MATLAB with MATPOWER [6.10] incorporated for the load flow.

6.5.1. Case 6.1: Univariate Non-Gaussian Random Variable Problem

In this section, the performance of the gUT technique in treating problems with non-Gaussian random uncertainties such as the wind power distribution is evaluated. The 6 bus test system (Appendix C) is modified to include a 55MW wind farm on bus 3 as described in Section 5.2.2 of Chapter 5. The case simply extends the results presented in Table 5.6 to include those for the gUT technique. The modified table is shown in Table 6.2.

From the table, it is seen that the gUT gives about the same level of accuracy as the 5 points conventional UT for the output mean values. Both methods outperform the other schemes presented. However, for higher order moments (e.g. standard deviation), gUT gives result better than those for the 5 points conventional UT and all the other techniques presented. This makes the gUT method a viable alternative to the cUT technique for problems with non-symmetrically distributed uncertainties.

Table 6.2: Average Percentage Error Indices for Case 6.1

Moments		3Pts cUT	5Pts cUT	PEM	CM	gUT
Mean ε_{μ}[%]	V	0.02057	8.081e-5	0.06756	0.02580	8.123e-5
	δ	0.31494	0.03421	1.05818	0.37482	0.03427
	P*i,j*	2.97800	0.27137	8.27511	3.85480	0.27156
	Q*i,j*	0.78000	0.00956	0.25623	0.97783	0.00958
Std ε_{σ}[%]	V	61.1553	0.11429	10.9163	14.0502	0.06013
	δ	54.9916	0.08715	5.78162	0.41455	0.08582
	P*i,j*	54.6800	0.08626	5.98060	0.54953	0.08611
	Q*i,j*	53.5722	0.09352	7.06495	19.6638	0.06973
Skewness $\varepsilon_{\gamma 1}$[%]	V	673.827	33.4753	232.885	191.312	8.86262
	δ	391.603	0.36885	18.0957	1.89796	0.19810
	P*i,j*	386.971	0.28718	18.6556	2.73878	0.21278
	Q*i,j*	459.132	27.9033	71.7575	139.917	9.35622
Kurtosis $\varepsilon_{\gamma 2}$[%]	V	1338.81	9.13140	38.9689	48.1847	12.3150
	δ	893.904	0.21952	6.26801	0.96885	0.43862
	P*i,j*	886.411	0.66233	6.89060	3.09220	0.39848
	Q*i,j*	968.017	8.48978	30.4196	37.7007	9.35003

6.5.2. Case 6.2: Multivariate Normally Distributed Variable Problem

In this section, the 6 bus test system is modified to include 2 normally distributed random variables; the active loads on buses 3 and 5. Both loads are assumed to be normally distributed with 10% coefficient of variation. Just as in the previous section, the results presented here extend those from Section 5.2.3. Table 6.3 presents the average percentage error results for the Taylor series based multivariate cUT (denoted T.S.M. in the table), the General Set cUT, the PEM, the CM and the gUT relative to the MCS method.

Table 6.3: Average Percentage Error Indices for Case 6.2

Moments		T.S.M	General Set	PEM	CM	gUT
Mean ε_{μ}[%]	V	0.00104	0.00104	0.00105	0.01070	0.00104
	δ	0.02679	0.02679	0.02687	0.10915	0.02679
	P*i,j*	0.03420	0.03419	0.03421	0.04457	0.03419
	Q*i,j*	0.03913	0.03912	0.03925	0.24871	0.03912
Std ε_{σ}[%]	V	0.06968	0.06968	0.12122	0.27158	0.06928
	δ	0.02936	0.02950	0.03705	0.09800	0.02947
	P*i,j*	0.06712	0.06704	0.06733	0.17777	0.06706
	Q*i,j*	0.04935	0.05255	0.39507	1.44371	0.04815
Skewness $\varepsilon_{\gamma 1}$[%]	V	2.25649	1.53661	43.4231	100.000	1.06202
	δ	5.23306	4.13746	58.1719	100.000	3.67126
	P*i,j*	10.2480	9.33190	81.7092	100.000	9.16827
	Q*i,j*	6.64280	3.42245	80.2789	100.000	2.72021
Kurtosis $\varepsilon_{\gamma 2}$[%]	V	2.70175	1.68679	30.2088	2.04714	0.55109
	δ	1.11309	0.58099	36.4028	0.64248	0.29551
	P*i,j*	0.83869	0.58995	30.2283	0.58742	0.56559
	Q*i,j*	6.94073	5.27384	22.3406	5.45092	0.33420

From the average error results, the gUT and the General Set cUT have equivalent accuracy in estimating the output mean values, and are more accurate than the other techniques. For the higher moment estimation, the gUT method outperforms all the other techniques as reflected in Table 6.3.

Table 6.4: Computation Time for Methods in Case 6.2

Method	MCS	T.S.M	General Set	PEM	CM	gUT
Number of Evaluations	1E5	6	9	5	1	25

From Table 6.4, the gUT method requires the highest estimation points aside from the MCS method. As already highlighted in Section 6.4, this high dependency between the number of random numbers and number of estimation

presents a major drawback to the gUT method and thus limits its applicability in systems with large number of uncertainties as will be later demonstrated.

6.5.3. Case 6.3: Multivariate Problem with Normal and Non-Normal Variables

The applicability of the gUT method to power systems with symmetrical and non-symmetrical multivariate problems is demonstrated in this section. The 6 bus test system is modified to include a 55MW rated wind farm on bus 3 and a normally distributed active load with 10% coefficient of variation located on bus 5. The wind speed and wind turbine parameters follow as those presented in Table 5.5 of Chapter 5.

Table 6.5: Results of Moments for Case 6.3 Showing Selected Values for the Mean and Active Power Flow

Moments		PEM	gUT	CM	MCS
Mean	V_3	1.27924	1.18159	1.18231	1.18159
	V_5	1.03970	0.96030	0.96042	0.96030
	$P_{1,6}$	-0.26558	-0.24515	-0.24503	-0.24515
	$P_{3,4}$	-0.18771	-0.17386	-0.17407	-0.17383
Std	V_3	0.10173	0.00413	0.00464	0.00413
	V_5	0.08273	0.00420	0.00413	0.00422
	$P_{1,6}$	0.04780	0.04080	0.04089	0.04073
	$P_{3,4}$	0.03631	0.02721	0.02740	0.02727
Skewness	V_3	-0.96343	0.17888	0.94960	0.17741
	V_5	-0.96517	-0.51855	-0.94960	-0.51831
	$P_{1,6}$	1.72803	0.74732	0.75406	0.75069
	$P_{3,4}$	-0.79536	-0.93502	-0.94041	-0.93697
Kurtosis	V_3	0.92968	2.13025	3.25602	2.25043
	V_5	0.93424	3.21211	3.11337	3.19427
	$P_{1,6}$	3.55840	3.21341	3.22897	3.21617
	$P_{3,4}$	2.59827	3.29404	3.30737	3.29315

A few selected results of the voltage and active power flow using the Gaussian quadrature based UT (gUT), the point estimate method (PEM), the cumulant method (CM) and the Monte Carlo simulation (MCS) method are presented in Table 6.5. The results for conventional UT method have been left out due to its resulting in abnormal (negative) sigma point and weight as detailed in Section 5.2.4.

Table 6.6: Average Percentage Error Indices for Case 6.3

Moments		PEM	CM	gUT
Mean	V	0.02819	8.26607	3.8136e-5
	δ	0.44385	8.50785	0.00965
	P*i,j*	5.07381	7.35756	0.09052
	Q*i,j*	1.10095	8.17229	0.00250
Std	V	6.84741	2342.02	0.22407
	δ	0.28884	9.36526	0.12621
	P*i,j*	0.47116	34.6031	0.20273
	Q*i,j*	18.4141	506.028	0.09623
Skewness	V	391.745	448.170	1.00866
	δ	2.71261	81.4243	0.42245
	P*i,j*	5.11503	270.293	0.25791
	Q*i,j*	305.231	482.106	8.04780
Kurtosis	V	21.6547	64.7261	2.57811
	δ	1.08476	12.4646	0.09580
	P*i,j*	0.75823	30.1477	0.30936
	Q*i,j*	31.5790	48.1251	6.71745

From the result, it is seen that the gUT method gives the closest result to the MCS method which is used as the benchmark. To further reflect the performance of the methods, the average percentage error indices up to the 4^{th} moments are shown in Table 6.6. The average error in the mean using the gUT method is less than 0.1% while it is as high as 5% and 8% in the cumulant method and the point estimate method respectively. For higher order moments (up to the 4^{th} moment), the highest average percentage error for the gUT is

about 8% while it is as high as 2000% and 500% in PEM and CM respectively. Overall, it is seen that the errors in the gUT method are much lower as compared with the other techniques. This shows the better performance of the gUT method for problems with non-symmetrical distributions, however the fundamental problem of *curse of dimensionality* still affects the number of estimation required as the number of uncertain parameters increases.

6.5.4. Case 6.4: Application in a Larger System

In this section the performance of the gUT method in a larger power system is explored in order to verify the scalability and accuracy of the method. To do this, the 24 bus IEEE reliability test system (Appendix C) is modified to include a wind farm and uncertainties in the load. The active and reactive loads on buses 4 and 8 are assumed to follow the normal distribution with a 5% coefficient of variation while the output wind power is fed into the system as a negative load on bus 17. The wind farm is made to supply 10% (285MW) of the total system active load. The wind farm is assumed to be an aggregate of the output from a group of 1MW wind turbines whose parameters and wind speed data are the same as those in Table 5.5.

The result for the average percentage error indices for the gUT, CM and PEM relative to the MCS method are presented in Table 6.7. From the result, the average error in the output mean is less than 0.02% for the gUT method while it is as high as 0.8% and 4.7% for PEM and CM. The average percentage error in the higher moments (up to the 4^{th} moment) is less than 2% for the gUT method. This shows the accuracy of the method relative to the MCS method. Overall, as reflected from the table, the gUT method gives the closest result to the MCS of all the techniques considered.

Comparing the results in Table 6.7 with those in Table 6.6, it is seen that the magnitude of the errors in the smaller system (with high wind penetration level) is higher. For case 3 above, the wind farm is rated about 60% of the total system load while in the current study, the wind farm supplies 10% of the total system load. From this, it is clear that the gUT method can successfully cope

with power systems with both small and large wind penetration while maintaining a high level of accuracy.

Table 6.7: Average Percentage Error Indices for Case 4

Moments		PEM	CM	gUT
Mean	V	7.16e-04	0.00279	7.08e-06
	δ	0.25937	0.07346	0.00366
	P*i,j*	0.25602	0.05827	0.00315
	Q*i,j*	0.79820	4.71000	0.01942
Std	V	2.89462	7.29478	0.03710
	δ	5.04631	0.28860	0.04110
	P*i,j*	4.49904	0.26371	0.04650
	Q*i,j*	4.70060	6.73787	0.04960
Skewness	V	37.9788	24.0047	0.39282
	δ	17.8648	2.13306	0.05687
	P*i,j*	209.842	13.6903	1.89860
	Q*i,j*	54.8458	128.617	1.15102
Kurtosis	V	23.8003	10.3439	1.03173
	δ	13.5346	1.09056	0.03345
	P*i,j*	14.0255	6.43209	0.05644
	Q*i,j*	33.8408	16.8572	1.31143

Table 6.8: Computation Time for Methods in Case 4

Method	MCS	PEM	CM	gUT
Computation Time (sec)	862.49	0.6642	0.0369	29.030
Number of Evaluations	1E5	11	1	3125

The number of computations and the estimation time required for implementing each of the techniques used is shown in Table 6.8. From the table, it is seen that the gUT method though very accurate requires more computation time as compared with the PEM and CM. Using the gUT method,

the number of estimation points required become the same as the MCS method (for this scenario with 5 sigma points for each variable) once the number of uncertainties is about 7. This limits the applicability of the gUT method in realistic power systems with large number of uncertain parameters.

6.6. Précis

The Gaussian quadrature based Unscented Transform technique has been presented in this chapter as an improvement to the conventional Unscented Transform method which is a mathematical estimator in lieu of the Monte Carlo Simulation method. The mathematical basis of orthogonal polynomial on which the proposed technique is anchored has been expounded. The technique has been applied to power systems with single and multivariate variables which vary either symmetrically or non-symmetrically. The results obtained have been compared against those from the Monte Carlo Simulation method which is taken as the benchmark as well as with other existing techniques.

The performance of the Gaussian quadrature based UT has been evaluated using 4 case studies with and without wind farm output. From the results obtained, the technique is seen to give the closest values to the Monte Carlo Simulation method of all the other techniques considered. However, for realistic problems with a large number of uncertainties, the number of estimations required becomes impractical and almost impossible. Though the technique is capable of accurately coping with both symmetrically and non-symmetrically distributed random variables, the *curse of dimensionality* problem associated with it limits its application in the modern day power system which are characterised by high number of uncertain parameters.

Considering that the ideal alternative technique is one which gives a good balance between speed and accuracy, a means of reducing the number of estimation required by the Gaussian quadrature based UT method while still maintaining its high level of accuracy is desired. One possible way of doing this, is using dimension reduction techniques which will be discussed in the next chapter.

6.7. References

[6.1] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in C*, 2nd ed., New York: Cambridge University Press, 1997.

[6.2] E. W. Weisstein, "Gaussian Quadrature," *MathWorld*-A Wolfram Web Resource, [Online]. Available: http://mathworld.wolfram.com/GaussianQuadrature.html [accessed June 2012]

[6.3] S. Rahman, "Extended polynomial dimensional decomposition for arbitrary probability distributions," *ASCE Journal of Engineering Mechanics*, pp. 1439-1451, December 2009.

[6.4] W. Gautschi, *Orthogonal Polynomials: computation and approximation*, New York: Oxford University press, 2004.

[6.5] G. Dalquist and A. Bjorck, *Numerical Methods in Scientific Computing*, vol. I, Society for Industrial and Applied Mathematics, 2007.

[6.6] A. D. Fernandes and W. R. Atchley, "Gaussian quadrature formulae for arbitrary positive measures," [Online]. Available: http://www.la-press.com/gaussian-quadrature-formulae-for-arbitrary-positive-measures-article-a134. [accessed Feb 2010]

[6.7] W. Gautschi, *Computational aspect of Orthogonal Polynomials* in, *Orthogonal Polynomial: Theory and Practice*, The Netherlands: Kluwer Academic Publishers, 1990.

[6.8] "Quadrature," *Digital Library of Mathematical Functions*, National Institute of Standards and Technology [Online]. Available: http://dlmf.nist.gov/. [accessed June 2012]

[6.9] M. S. Eldred, "Design under uncertainty employing stochastic expansion methods," *Inter Journal for Uncertainty Quantification*, vol. 1, no. 2, pp. 119-146, 2011.

[6.10] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER's Extensible Optimal Power Flow Architecture," *Power and Energy Society General Meeting, 2009 IEEE*, pp. 1-7, July 26-30 2009.

[6.11] P. J. Davis, and P. Rabinowitz, *Methods of Numerical Integration*, New York: Academic Press Inc., 1975.

Chapter 7

Enhanced Unscented Transform Method

The *curse of dimensionality* problem has been identified as the main limitation to applying the Unscented Transform method in practical systems with large number of uncertainties. In this chapter, a dimension reduction technique is combined with the UT method to drastically reduce the number of required evaluations. The mathematical basis of the technique is considered while the improvement achievable by using it is demonstrated using both transmission and distribution power systems.

7.1. Mathematical Basics of Dimension Reduction

Dimension reduction [7.1-7.3] is a technique for function approximation to estimate the statistical moments of the output function. The technique involves an additive decomposition of an n-dimensional function involving n-dimensional integral into a series sum of D-dimensional functions such that $D<n$. Simply put, the dimension reduction provides a way of efficiently combining the sigma points and weights for a large number of variable such that the number of evaluation points can be minimised. For the case where D=1, the technique is referred to as Univariate dimension reduction while it is referred to as Bivariate when D=2. The principle behind the decomposition is described in the next section.

7.1.1. Univariate Dimension Reduction (UDR)

With the UDR, the main function is decomposed into a series of n-dimensional function thus reducing the problem to a one-dimensional integration; it can be employed in problems involving a minimum of two random variables. The mathematical basis is described below.

For a function $f(\boldsymbol{x})$ with N independent variables such that;

$$y = f(\mathbf{x}) \tag{7.1}$$

$$\mathbf{x} = \{x_1, x_2, \ldots, x_N\}$$

From (6.17), the mean of the function can be written as;

$$E[f(\mathbf{x})] = \int_a^b \cdots \int_a^b f(\mathbf{x}) d\mathbf{x} \tag{7.2}$$

where *a* and *b* depends on the type of distribution.

This involves *N*-dimensional integration of (7.2).

At the mean points, let;

$$\mathbf{x}_\mu = (\bar{\mathbf{x}}) = \{\bar{x}_1, \ldots, \bar{x}_N\}^T \tag{7.3}$$

Expanding *f(x) using* Taylor series expansion gives;

$$y = f(\mathbf{x}) = f(\bar{\mathbf{x}} + \hat{\mathbf{x}}) = f(\bar{\mathbf{x}}) + \sum_{j=1}^{\infty} \frac{1}{j!} \sum_{i=1}^{N} (\bar{\mathbf{x}}) \frac{d^j f}{d\hat{x}_i^j} \hat{x}_i^j + \sum_{j_2=1}^{\infty} \sum_{j_1=1}^{\infty} \frac{1}{j_1! j_2!} \sum_{i1<i2}^{N} (\bar{\mathbf{x}}) \frac{d^{j_1+j_2} f}{d\hat{x}_{i_1}^{j_1} d\hat{x}_{i_2}^{j_2}} \hat{x}_{i_1}^{j1} \hat{x}_{i_2}^{j2}$$
$$+ \sum_{j_3=1}^{\infty} \sum_{j_2=1}^{\infty} \sum_{j_1=1}^{\infty} \frac{1}{j_1! j_2!} \sum_{i_1<i_2}^{N} (\bar{\mathbf{x}}) \frac{d^{j_1+j_2+j_3} f}{d\hat{x}_{i_1}^{j_1} d\hat{x}_{i_2}^{j_2} d\hat{x}_{i_3}^{j_3}} \hat{x}_{i_1}^{j1} \hat{x}_{i_2}^{j2} \hat{x}_{i_3}^{j3} + \cdots \tag{7.4}$$

Substituting (7.4) into (7.2);

$$E[f(\mathbf{x})] = E[f(\bar{\mathbf{x}})] + \sum_{j=1}^{\infty} \frac{1}{j!} \sum_{i=1}^{N} (\bar{\mathbf{x}}) \frac{d^j f}{d\hat{x}_i^j} E[\hat{x}_i^j] + \sum_{j_2=1}^{\infty} \sum_{j_1=1}^{\infty} \frac{1}{j_1! j_2!} \sum_{i_1<i_2}^{N} (\bar{\mathbf{x}}) \frac{d^{j_1+j_2} f}{d\hat{x}_{i_1}^{j_1} d\hat{x}_{i_2}^{j_2}} E[\hat{x}_{i1}^{j1} \hat{x}_{i2}^{j2}]$$
$$+ \sum_{j_3=1}^{\infty} \sum_{j_2=1}^{\infty} \sum_{j_1=1}^{\infty} \frac{1}{j_1! j_2!} \sum_{i_1<i_2}^{N} (\bar{\mathbf{x}}) \frac{d^{j_1+j_2+j_3} f}{d\hat{x}_{i_1}^{j_1} d\hat{x}_{i_2}^{j_2} d\hat{x}_{i_3}^{j_3}} E[\hat{x}_{i_1}^{j1} \hat{x}_{i_2}^{j2} \hat{x}_{i_3}^{j3}] + \cdots \tag{7.5}$$

For a univariate function $f(\bar{x}_1, \ldots, x_i, \bar{x}_{i+1}, \ldots, \bar{\mu}_N)$, the Taylor series expansion is given by (7.6).

$$f(\bar{x}_1, \ldots, x_i, \bar{x}_{i+1}, \ldots, \bar{x}_N) = f(\bar{\mathbf{x}}) + \sum_{j=1}^{\infty} \frac{1}{j!} \sum_{i=1}^{N} (\bar{\mathbf{x}}) \frac{d^j f}{dx_i^j} \hat{x}_i^j \tag{7.6}$$

Using (7.6), the Taylor series expansion of the first N univariate functions is given as;

$$\begin{aligned} f(x_1, \bar{x}_2, \bar{x}_3, \ldots, \bar{x}_N) &= f(\bar{\mathbf{x}}) + \sum_{j=1}^{\infty} \frac{1}{j!} (\bar{\mathbf{x}}) \frac{d^j f}{dx_1^j} \hat{x}_1^j \\ f(\bar{x}_1, x_2, \bar{x}_3, \ldots, \bar{x}_N) &= f(\bar{\mathbf{x}}) + \sum_{j=1}^{\infty} \frac{1}{j!} (\bar{\mathbf{x}}) \frac{d^j f}{dx_2^j} \hat{x}_2^j \\ &\vdots \\ f(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{N-1}, x_N) &= f(\bar{\mathbf{x}}) + \sum_{j=1}^{\infty} \frac{1}{j!} (\bar{\mathbf{x}}) \frac{d^j f}{dx_N^j} \hat{x}_N^j \end{aligned} \tag{7.7}$$

Summing these single functions in (7.7) gives

$$\sum_{i=1}^{N} f(\bar{x}_1, \bar{x}_2, \ldots, x_i, \bar{x}_N) = N f(\bar{\mathbf{x}}) + \sum_{i=1}^{N} \sum_{j=1}^{\infty} \frac{1}{j!} (\bar{\mathbf{x}}) \frac{d^j f}{dx_i^j} \hat{x}_i^j \tag{7.8}$$

If $\hat{f}(\mathbf{x})$ is a univariate approximation of *f(x)* which is given by;

$$\hat{f}(\mathbf{x}) = \sum_{i=1}^{N} f(\bar{x}_1, \bar{x}_2, \ldots, x_i, \bar{x}_N) - (N-1) f(\bar{\mathbf{x}}) \tag{7.9}$$

The mean of the univariate approximation $\hat{f}(\mathbf{x})$ is;

$$E[\hat{f}(\mathbf{x})] = \sum_{i=1}^{N} E[f(\bar{x}_1, \bar{x}_2, \ldots, x_i, \bar{x}_N)] - (N-1) E[f(\bar{\mathbf{x}})] \tag{7.10}$$

Substituting (7.8) into (7.10) gives;

$$E[\hat{f}(\mathbf{x})] = E[f(\bar{\mathbf{x}})] + \sum_{i=1}^{N} \sum_{j=1}^{\infty} \frac{1}{j!} (\bar{\mathbf{x}}) \frac{d^j f}{d\hat{x}_i^j} \hat{x}_i^j \tag{7.11}$$

Assuming the higher order terms are neglected, $E[\hat{f}(\mathbf{x})]$ and $E[f(\mathbf{x})]$ are approximately equal to each other. Therefore, the output moments of the function y are easily assumed as $E[\hat{f}(\mathbf{x})]$ which is a summation of one-dimensional functions and involves only one dimensional integration. The residual error between (7.5) and (7.11) denoted as eu for the UDR is;

$$e_u = E[f(\mathbf{x})] - E[\hat{f}(\mathbf{x})] = \sum_{j_2=1}^{\infty}\sum_{j_1=1}^{\infty}\frac{1}{j_1!\,j_2!}\sum_{i_1<i_2}^{N}(\bar{\mathbf{x}})\frac{d^{\,j_1+j_2}f}{d\hat{x}_{i_1}^{\,j_1}d\hat{x}_{i_2}^{\,j_2}}E[\hat{x}_{i_1}^{\,j1}\hat{x}_{i_2}^{\,j2}] + \sum_{j_3=1}^{\infty}\sum_{j_2=1}^{\infty}\sum_{j_1=1}^{\infty}\frac{1}{j_1!\,j_2!}\sum_{i_1<i_2}^{N}(\bar{\mathbf{x}})\frac{d^{\,j_1+j_2+j_3}f}{d\hat{x}_{i_1}^{\,j_1}d\hat{x}_{i_2}^{\,j_2}d\hat{x}_{i_3}^{\,j_3}}E[\hat{x}_{i_1}^{\,j1}\hat{x}_{i_2}^{\,j2}\hat{x}_{i_3}^{\,j3}] + \cdots \quad (7.12)$$

The error e_u in the univariate function can become significant for systems with large random variation [7.1] and for higher order moments. To reduce the residual error in (7.12), the bivariate dimension reduction can be employed although, at the expense of computational time.

7.1.2. Bivariate Dimension Reduction (BDR)

The bivariate dimension reduction provides an improvement to the univariate reduction discussed above since only the third and higher order terms (which are most times trivial) are now contained in its residual error (e_b). With the BDR, a combination of one and two-dimensional functions are used in approximating the function. The BDR can be applied to problems involving three or more random variables.

Consider a bivariate function $f(\bar{x}_1,\ldots,x_{i1},\bar{x}_{i+1},\ldots,x_{i2},\ldots,\bar{x}_N)$ with its Taylor series expansion shown in (7.13);

$$f(\bar{x}_1,\ldots,x_{i1},\bar{x}_{i+1},\ldots,x_{i2},\ldots,\bar{x}_N) = f(\bar{\mathbf{x}}) + \sum_{j=1}^{\infty}\frac{1}{j!}\sum_{i=1}^{N}(\bar{\mathbf{x}})\frac{d^{\,j}f}{d\hat{x}_{i1}^{\,j}}\hat{x}_{i1}^{\,j} + \sum_{j=1}^{\infty}\frac{1}{j!}\sum_{i=1}^{N}(\bar{\mathbf{x}})\frac{d^{\,j}f}{d\hat{x}_{i2}^{\,j}}\hat{x}_{i2}^{\,j} + \sum_{j_2=1}^{\infty}\sum_{j_1=1}^{\infty}\frac{1}{j_1!\,j_2!}\sum_{i_1<i_2}^{N}(\bar{\mathbf{x}})\frac{d^{\,j_1+j_2}f}{d\hat{x}_{i_1}^{\,j_1}d\hat{x}_{i_2}^{\,j_2}}\hat{x}_{i_1}^{\,j1}\hat{x}_{i_2}^{\,j2} \quad (7.13)$$

Using (7.13), the summation of the one and two-dimensional functions for a function made up of N variables is given by [7.3];

$$\sum_{i1<i2}^{N} f(\bar{x}_1,\ldots,x_{i1},\bar{x}_{i+1},\ldots,x_{i2},\ldots,\bar{x}_N) = \frac{N(N-1)}{2}f(\bar{\mathbf{x}}) + (N-1)\sum_{i=1}^{N}\sum_{j=1}^{\infty}\frac{1}{j!}(\bar{\mathbf{x}})\frac{d^{\,j}f}{d\hat{x}_i^{\,j}}\hat{x}_i^{\,j} + \sum_{i_1<i_2}\sum_{j_2=1}^{\infty}\sum_{j_1=1}^{\infty}\frac{1}{j_1!\,j_2!}\sum_{i_1<i_2}^{N}(\bar{\mathbf{x}})\frac{d^{\,j_1+j_2}f}{d\hat{x}_{i_1}^{\,j_1}d\hat{x}_{i_2}^{\,j_2}}\hat{x}_{i_1}^{\,j1}\hat{x}_{i_2}^{\,j2} \quad (7.14)$$

The bivariate approximation of the Taylor series expansion of *f(x)* in (7.4) can be written in terms of $\sum_{i1<i2}^{N} f(\bar{x}_1,\ldots,x_{i1},\bar{x}_{i+1},\ldots,x_{i2},\ldots,\bar{x}_N)$ and $\sum_{i=1}^{N} f(\bar{x}_1,\bar{x}_2,\ldots,x_i,\bar{x}_N)$ as;

$$\hat{f}(\mathbf{x}) = \sum_{i1<i2}^{N} f(\bar{x}_1,\ldots,x_{i1},\bar{x}_{i+1},\ldots,x_{i2},\ldots,\bar{x}_N)\cdots - (N-2)\sum_{i=1}^{N} f(\bar{x}_1,\bar{x}_2,\ldots,x_i,\bar{x}_N) + \frac{(N-1)(N-2)}{2} f(\bar{\mathbf{x}}) \tag{7.15}$$

The mean of $\hat{f}(\mathbf{x})$ is given as;

$$E[\hat{f}(\mathbf{x})] = \sum_{i1<i2}^{N} E[f(\bar{x}_1,\ldots,x_{i1},\bar{x}_{i+1},\ldots,x_{i2},\ldots,\bar{x}_N)] - (N-2)\sum_{i=1}^{N} E[f(\bar{x}_1,\bar{x}_2,\ldots,x_i,\bar{x}_N)] + \frac{(N-1)(N-2)}{2} E[f(\bar{\mathbf{x}})] \tag{7.16}$$

The difference between (7.5) and (7.16) gives the residual error e_b (7.17) in the bivariate approximation. As expected, the bivariate approximation gives better results than the univariate approximation especially for the higher other moments since it includes more higher order terms in the approximation.

$$e_b = \sum_{j_3=1}^{\infty}\sum_{j_2=1}^{\infty}\sum_{j_1=1}^{\infty}\frac{1}{j_1!\,j_2!}\sum_{i_1<i_2}^{N}(\bar{\mathbf{x}})\frac{d^{\,j_1+j_2+j_3} f}{d\hat{x}_{i_1}^{j_1} d\hat{x}_{i_2}^{j_2} d\hat{x}_{i_3}^{j_3}} E[\hat{x}_{i_1}^{j1}\hat{x}_{i_2}^{j2}\hat{x}_{i_3}^{j3}] + \cdots \tag{7.17}$$

7.1.3. Output Moment Estimation

The higher order moment of the output function can easily be estimated using the dimension reduction technique. The *nth* moment $E[f^n(x)]$ is simply evaluated using a simple assumption [7.1]. For instance, the *nth* order of the function *f(x)* can be represented by *Z(x)* as shown in (7.18) for easy expansion;

$$f^n(x) = Z(x) \tag{7.18}$$

For the UDR, the higher moment is given by (7.19);

$$E[f^n(\mathbf{x})] = E[Z(\mathbf{x})] \cong E\left\{\sum Z(\bar{x}_1,\ldots,x_i,\bar{x}_{i+1},\ldots,\bar{x}_N) - (N-1)Z(\bar{\mathbf{x}})\right\} \tag{7.19}$$

Substituting (7.18) back into the right side of (7.19);

$$E[f^n(\mathbf{x})] = E[Z(\mathbf{x})] \cong E\left\{\sum f^n(\bar{x}_1,\ldots,x_i,\bar{x}_{i+1},\ldots,\bar{x}_N) - (N-1)f^n(\bar{\mathbf{x}})\right\} \quad (7.20)$$

Following the same principle the central moment about the mean μ_y is given by;

$$E[f^n(\mathbf{x}-\mu_y)] \cong E\begin{Bmatrix}\sum(f^n(\bar{x}_1,\ldots,x_i,\bar{x}_{i+1},\ldots,\bar{x}_N)-\mu_y)\\ -(N-1)(f^n(\bar{\mathbf{x}})-\mu_y)\end{Bmatrix} \quad (7.21)$$

Equally, the raw and central moments of the output function using the BDR technique are given in (7.22) and (7.23) respectively.

$$E[f^n(\mathbf{x})] \cong E\begin{bmatrix}\sum_{i1<i2}^{N} f^n(\bar{x}_1,\ldots,x_{i1},\bar{x}_{i_1+1},\ldots,x_{i2},\ldots,\bar{x}_N)\\ -(N-2)\sum_{i=1}^{N} f^n(\bar{x}_1,\ldots,x_i,\bar{x}_{i+1},\ldots,\bar{x}_N)\\ +\dfrac{(N-1)(N-2)}{2} f^n(\bar{\mathbf{x}})\end{bmatrix} \quad (7.22)$$

$$E[f^n(\mathbf{x}-\mu_y)] \cong E\begin{bmatrix}\sum_{i1<i2}^{N}(f^n(\bar{x}_1,\ldots,x_{i1},\bar{x}_{i_1+1},\ldots,x_{i2},\ldots,\bar{x}_N)-\mu_y)\\ -(N-2)\sum_{i=1}^{N}(f^n(\bar{x}_1,\ldots,x_i,\bar{x}_{i+1},\ldots,\bar{x}_N)-\mu_y)\\ +\dfrac{(N-1)(N-2)}{2}(f^n(\bar{\mathbf{x}})-\mu_y)\end{bmatrix} \quad (7.23)$$

7.2. Dimension Reduction and Unscented Transform

The dimension reduction technique is applied to the UT method based on the relationship between the moments of a continuous function and its discrete approximation (7.24).

$$E[f(x)] = \sum_{i=1}^{N} w_i f(S_i) \quad (7.24)$$

where w_i are the weights and S_i the sigma points.

Substituting the values of *f(x)* into (7.24) gives (7.25) for the UDR techniques.

$$E[f(\mathbf{x}) \cong E[\hat{f}(\mathbf{x})] \cong \sum_{i=1}^{N} E[f(\bar{x}_1, \bar{x}_2, \ldots, x_i, \ldots, \bar{x}_N)] - (N-1)E[f(\bar{x}_1, \ldots, \bar{x}_N)]$$

$$\cong \sum_{j=1}^{n}\sum_{i=1}^{N} w_i f(\bar{x}_1, \bar{x}_2, \ldots, S_i^j, \ldots, \bar{x}_N) - (N-1) f(\bar{x}_1, \ldots, \bar{x}_N) \quad (7.25)$$

The higher order raw moment for the UDR can be evaluated using (7.26).

$$E[f^n(\mathbf{x})] \cong \sum_{j=1}^{n}\sum_{i=1}^{N} w_i f^n(\bar{x}_1, \bar{x}_2, \ldots, S_i^j, \bar{x}_{i+1}, \ldots, \bar{x}_N) - (N-1) f^n(\bar{x}_1, \ldots, \bar{x}_N) \quad (7.26)$$

The central moment is obtained following the same principle as;

$$E[f^n(\mathbf{x} - \mu_y)] \cong \sum_{j=1}^{n}\sum_{i=1}^{N} w_i (f^n(\bar{x}_1, \bar{x}_2, \ldots, S_i^j, \bar{x}_{i+1}, \ldots, \bar{x}_N) - \mu_y) - (N-1)(f^n(\bar{x}_1, \ldots, \bar{x}_N) - \mu_y) \quad (7.27)$$

The raw and central higher moments using the BDR technique are given in (7.28) and (7.29) respectively.

$$E[f^n(\mathbf{x})] \cong \sum_{j=1}^{n}\sum_{i1<i2}^{N} w_i^j f^n(\bar{x}_1, \ldots, S_{i_1}^j, \bar{x}_{i_1+1}, \ldots, S_{i_2}^j, \bar{x}_N) - (N-2)\sum_{j=1}^{n}\sum_{i=1}^{N} w_i^j f^n(\bar{x}_1, \ldots, S_i^j, \bar{x}_{i+1}, \ldots, \bar{x}_N) + \frac{(N-1)(N-2)}{2} f^n(\bar{x}_1, \ldots, \bar{x}_N) \quad (7.28)$$

$$E[f(\mathbf{x} - \mu_y)^n] \cong \sum_{j=1}^{n}\sum_{i1<i2}^{N} w_i^j (f^n(\bar{x}_1, \ldots, S_{i_1}^j, \bar{x}_{i_1+1}, \ldots, S_{i_2}^j, \bar{x}_N) - \mu_y) - (N-2)\sum_{j=1}^{n}\sum_{i=1}^{N} w_i^j (f^n(\bar{x}_1, \ldots, S_i^j, \bar{x}_{i+1}, \ldots, \bar{x}_N) - \mu_y) + \frac{(N-1)(N-2)}{2}(f^n(\bar{x}_1, \ldots, \bar{x}_N) - \mu_y) \quad (7.29)$$

The number of evaluations (N_{EV}) required for an N-dimensional function using n estimation (sigma points) is given by (7.30) and (7.31) for the UDR and BDR techniques respectively [7.2];

$$N_{EV} = (n \times N) + 1 \quad (7.30)$$

$$N_{EV} = \frac{N(N-1)}{2} \times n^2 + ((n \times N) + 1) \tag{7.31}$$

In problems where all random variables are symmetrical and the number of sigma points *n* is odd, one of the sigma points will always be located on the mean point. Since the last set of estimation using the UDR is done at the mean point (see 7.10), the extra (redundant) point can be removed. For identification, this will be referred to as the reduced UDR and denoted as UDR~. The number of estimation required for the reduced UDR can be estimated using;

$$N_{EV} = ((n-1) \times N) + 1 \tag{7.32}$$

For N_{RV} random variables, the number of evaluation point required using the tensor product (section 6.4, Chapter 6) gUT, the univariate dimension reduction based UT (UT+UDR), the bivariate reduction technique UT (UT+BDR) and the reduced univariate dimension reduction (UT+UDR~) are shown in Table 7.1.

Table 7.1: Number of Evaluation Points Required for Using Various Methods (3 Point Approximation)

N_{RV}	gUT (Chapter 6)	UT+ UDR	UT+ BDR	UT+ UDR~
2	9	7	-	5
3	27	10	37	7
4	81	13	67	9
5	243	16	106	11
10	59049	31	436	21
20	3.49E+09	61	1771	41
30	2.06E+14	91	4006	61
40	1.22E+19	121	7141	81
50	7.18E+23	151	11176	101
100	5.15E+47	301	44851	201

From Table 7.1, it is clear that the gUT method implemented using the full dimensional integration becomes intractable when large number of random variables is involved. For instance, if 100,000 simulation points are assumed for the MCS method, only 10 random variables can be represented by the gUT method to have a smaller number of evaluation points than the MCS. While the UT plus univariate or bivariate dimension reduction method can accommodate up to 33,333 and 150 random variables respectively as compared to the MCS method. More random variables can be accommodated for cases where all the random variables are symmetrical. In this case, the UT+UDR~ can handle close to 50,000 random variables conveniently when compared with the MCS method (assumed to require 100,000 samples).

7.3. Case Studies

This section focuses on the performance appraisal of the dimension reduction based UT. Considering that the dimension reduction technique was proposed with the aim of reducing the number of evaluation points associated with the UT method while treating multivariate problems, its accuracy in solving such problems will be evaluated. To achieve this, three case studies are considered. In the first two cases, the technique is applied to a transmission test system with 3 and 24 random variables respectively while in the latter scenario, its applicability in a three phase unbalanced real distribution system power flow is demonstrated to show the versatility of the technique.

7.3.1. Case Study 7.1: Multivariate System with 3 Random Uncertainties

This case is aimed at evaluating the performance of the univariate and bivariate dimension reduction method as compared with the unscented transform (from Chapter 6). The results are also compared with those from the Monte Carlo Simulation (MCS) method (the benchmark), the point estimate method (PEM) and the Cumulant method (CM).

For the study, the 14-bus IEEE test system (Appendix C) is modified to include a 50MW rated wind farm on bus 6 while the active and reactive load demands on bus 9 are assumed to have 5% coefficient of variation. The wind turbine parameter and wind speed data are the same as those presented in Table 5.5 (Chapter 5). Two different approximation orders have been employed for the UDR method to understand the relationship between the number of sigma points and accuracy. A few selected results are shown in Table 7.2.

Table 7.2: Moments for the IEEE 14-Bus Test System Showing Selected Results

Moments		UT (5 pts.)	UT+UDR (5 pts.)	UT+UDR (9 pts.)	UT+BDR (5 pts.)	PEM	CM	MCS
Mean	V_{11}	1.05549	1.05549	1.05549	1.05549	1.05547	1.05555	1.05549
	δ_{11}	-11.8004	-11.8004	-11.8004	-11.8004	-11.7789	-11.8452	-11.8008
	$P_{6\text{-}11}$	-0.12422	-0.12422	-0.12422	-0.12422	-0.12460	-0.12339	-0.12422
	$Q_{6\text{-}11}$	-0.00164	-0.00164	-0.00164	-0.00164	-0.00163	-0.00166	-0.00164
Std	V_{11}	0.00086	0.00087	0.00087	0.00086	0.00091	0.00082	0.00086
	δ_{11}	1.35028	1.35324	1.35324	1.35028	1.42698	1.3545	1.34975
	$P_{6\text{-}11}$	0.02332	0.02338	0.02338	0.02332	0.02659	0.02334	0.02331
	$Q_{6\text{-}11}$	0.00080	0.00080	0.00080	0.00080	0.00080	0.00080	0.00080
Skewness	V_{11}	-0.77562	-0.69000	-0.69007	-0.77562	-0.69754	-0.60136	-0.77583
	δ_{11}	0.91851	0.91153	0.91153	0.91851	1.08534	0.93226	0.91664
	$P_{6\text{-}11}$	-0.92823	-0.91966	-0.91966	-0.92823	-1.0909	-0.93152	-0.92655
	$Q_{6\text{-}11}$	0.80173	0.79809	0.79806	0.80173	0.95748	0.87533	0.79962
Kurtosis	V_{11}	3.51020	2.24850	2.24317	3.50989	2.13456	3.16934	3.51947
	δ_{11}	3.27082	3.17072	3.17115	3.27082	3.02298	3.30382	3.26751
	$P_{6\text{-}11}$	3.29561	3.18773	3.18773	3.29561	3.0255	3.3035	3.29159
	$Q_{6\text{-}11}$	3.11047	2.79309	2.79522	3.11047	2.7847	3.2793	3.10920
No of Load flow Evaluations		125	16	28	91	7	1	100000

From the results shown in Table 7.2 and Table 7.3, it can be seen that the UT method (using full tensor product) and the UT+BDR technique give

approximately the same results. These (UT and UT+BDR) are the closest of all the techniques to the Monte Carlo Simulation method. The performance of the univariate dimension reduction based UT technique deteriorated when used for evaluating higher order moments. This is due to its neglecting the higher order terms in the Taylor series expansion as seen in (7.12).

Table 7.3: Average Percentage Error for the Moments

Moments		UT (5 pts.)	UT+UDR (5 pts.)	UT+UDR (9 pts.)	UT+BDR	PEM	CM
Mean	V	2.890e-5	2.890e-5	2.897e-5	2.889e-5	8.936e-4	2.834e-3
	δ	0.0024	0.0024	0.0024	0.0024	0.13575	0.27187
	P*i,j*	0.00530	0.00530	0.00529	0.00530	1.46693	3.11810
	Q*i,j*	0.02353	0.02354	0.02354	0.02353	1.82387	3.44235
Std	V	0.01442	0.6764	0.67630	0.01442	2.47971	1.97109
	δ	0.03652	0.23134	0.23134	0.03652	5.26327	0.38728
	P*i,j*	0.03734	0.26339	0.26339	0.03734	5.3474	0.24090
	Q*i,j*	0.03911	0.44638	0.44641	0.03911	4.98341	2.56234
Skewness	V	1.31790	161.965	161.180	1.32252	168.684	18.6450
	δ	0.18941	0.54810	0.54863	0.18941	16.9088	1.87739
	P*i,j*	0.17406	1.03518	1.03534	0.17406	15.6973	1.13229
	Q*i,j*	0.22365	1.77200	1.75407	0.22366	18.8230	14.8032
Kurtosis	V	0.13628	49.6853	49.6721	0.15574	20.5371	18.9201
	δ	0.09209	3.7713	3.7601	0.09209	7.6448	1.20856
	P*i,j*	0.10183	6.43765	6.43412	0.10185	10.3272	0.71495
	Q*i,j*	0.24697	9.89394	9.65268	0.24729	11.5487	7.51313

From both tables, it is clear that increasing the number of sigma points for the Univariate dimension reduction did not provide any significant improvement on the higher order moments. Overall, the PEM gave the poorest results partly due to its poor representation of the wind power distribution. The Cumulant method (CM) gave a fair output considering that only one computation is required. However, its (CM) performance for problems involving distributions

with active points away from the mean (e.g. the wind power distribution) can be improved by linearizing the variable at more than one point. This linearization is explained in [7.4-7.5]. As shown in the results, the Bivariate dimension reduction based UT stands as the best alternative to the MCS method, since it requires less evaluation points to achieve the same level of accuracy as the UT method.

7.3.2. Case Study 7.2: Multivariate Problem with 24 Random Variables

In Case 7.2, the performance of the techniques in coping adequately with systems with large number of random uncertainties is examined. The IEEE 14-bus test system is modified to include 24 uncertain variables that comprise the forced outage rate of the generators on buses 1 and 2 (0.9, 0.91 respectively) which follow the discrete distribution, while the 50MW wind farm is maintained on bus 6. The demand on all buses with non-zero active and reactive load was assumed to follow the normal distribution with a 5% coefficient of variation.

The percentage error using the UT based univariate and bivariate dimension reduction are presented in Table 7.4 alongside with those for the point estimate method and the cumulant method. All the results have been compared against those for Monte Carlo simulation with 100,000 samples. Results for the ordinary UT method have been left out due to its impracticality in treating the current scenario, as about 5.96×10^{16} evaluations are needed.

The result obtained affirms the adequacy of the UT based univariate dimension reduction for accurate estimation of the first and second moments only. The result for the PEM is seen to conform to those presented in a similar research in [7.6].

Table 7.4: Average Percentage Error for the Moments (Case 7.2)

Moments		UT+UDR (5 pts.)	UT+BDR	PEM	CM
Mean	V	1.828e-4	1.83e-4	9.719e-4	3.146e-3
	δ	0.04365	0.04365	0.17232	0.21305
	P*i,j*	0.50432	0.50430	2.79234	4.35931
	Q*i,j*	0.42107	0.42106	1.46252	1.11687
Std	V	0.12635	0.11123	1.81062	1.51817
	δ	0.08233	0.08296	3.95054	1.69866
	P*i,j*	0.07311	0.07393	4.37513	7.83025
	Q*i,j*	0.13095	0.13041	4.35830	13.5140
Skewness	V	61.1346	5.95484	76.6583	19.7305
	δ	2.0677	0.69019	17.0443	7.26355
	P*i,j*	11.2997	1.84479	118.111	144.053
	Q*i,j*	26.7272	2.06121	195.520	357.961
Kurtosis	V	33.7452	0.17912	33.0961	3.14476
	δ	21.4277	0.024111	20.4389	1.42109
	P*i,j*	24.1468	0.33187	25.0374	1.60607
	Q*i,j*	28.0840	0.3985	26.4476	6.00551

7.3.3. Case 7.3: A Practical Distribution System

Until now, the applicability of Unscented Transform method has only been tested on transmission systems. In this current scenario, the performance of the method on a real distribution system with unbalanced three phase load and embedded wind generation is examined.

The test system, which is a real power network in Samsun, Northern Turkey is made up of 44 buses. To ensure the scenario is close to reality as much as possible, the real measured wind data for the location, which is analysed in Appendix B, is used. A one-line diagram of the test system and the full details

of the test system including the Line, Load and Transformer parameters are presented in Appendix C.

Considering the 500kW limit for non-licenced wind farms in Turkey, it is assumed that the output from the 500kW rated wind farm is equally shared on all the three phases. The loads on all load buses with functioning loads were assumed to be uncertain and follow the normal distribution based on the average measured data from the site. In all, the network is made up of 117 uncertain random variables. The ladder network based backward/forward sweep load flow method described in Section 2.3.2.1 (Chapter 2) for distribution systems is employed.

The results obtained using the Univariate and Bivariate based dimension reduction techniques are compared with those for the Monte Carlo Simulation method in the following analysis. Comparison against previously used schemes (Point Estimate Method and Cumulant Method), have been omitted since peer reviewed publications detailing their proper application in distribution systems were not found.

The mean voltage magnitude for Phase A using the three techniques are presented in Fig. 7.1 while the mean percentage error in the voltage magnitude on all the buses using the Univariate and Bivariate dimension reduction UT techniques (relative to the Monte Carlo Simulation method) is illustrated in Figs. 7.2-7.4. From the plots, it is seen that both techniques give a close performance to the Monte Carlo Simulation method with the highest percentage error on all buses of all phases less than 4×10^{-4}. The system is also seen to conform with the stipulated voltage magnitude as all the voltages are within ±10% of the rated. To access the performance of the method in evaluating higher order moments (required for PDF and CDF curves), the standard deviation, skewness and kurtosis of the voltage magnitude for Phase A are shown in Figs. 7.5–7.7.

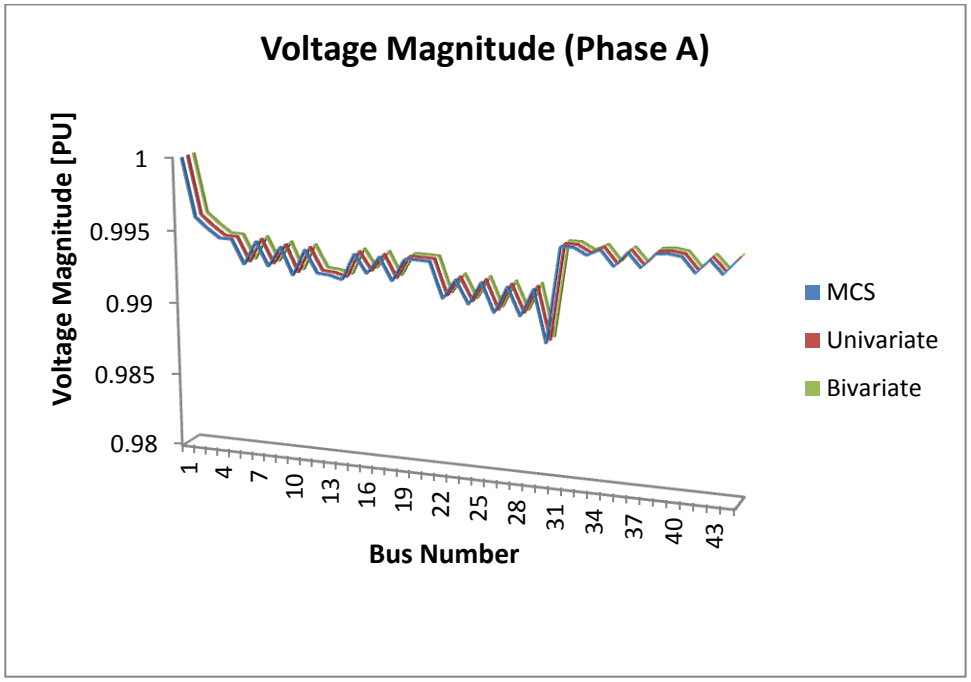

Fig 7.1: Mean Voltage Magnitude For Phase A Using the MCS, UT+UDR and UT+BDR Methods

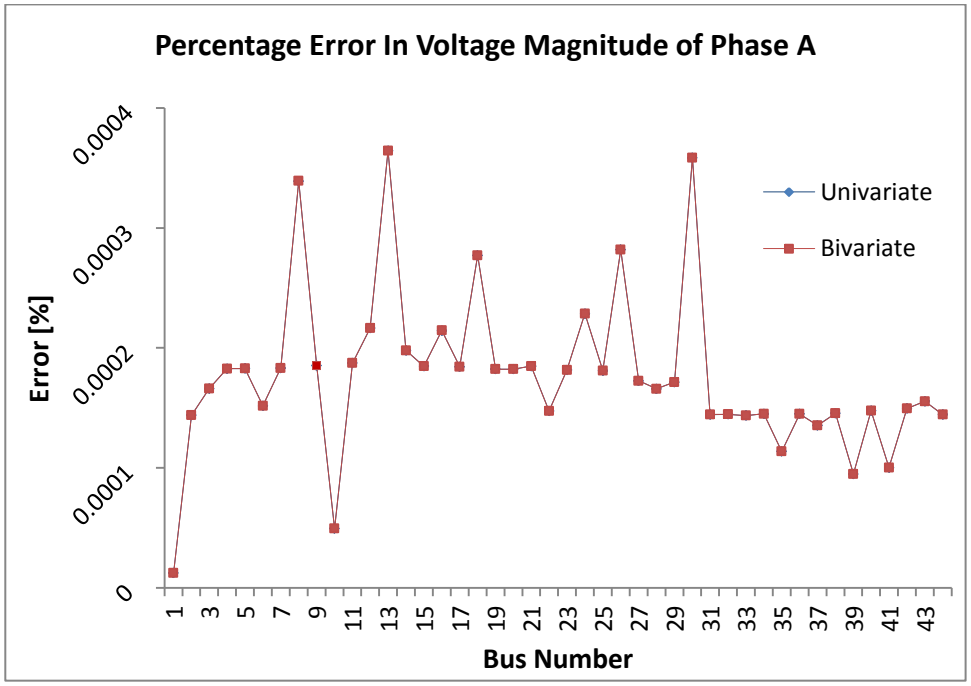

Fig 7.2: Average Percentage Error in Voltage Magnitude for Phase A Using the UT+UDR and UT+BDR Methods. (*Both give almost the same results*)

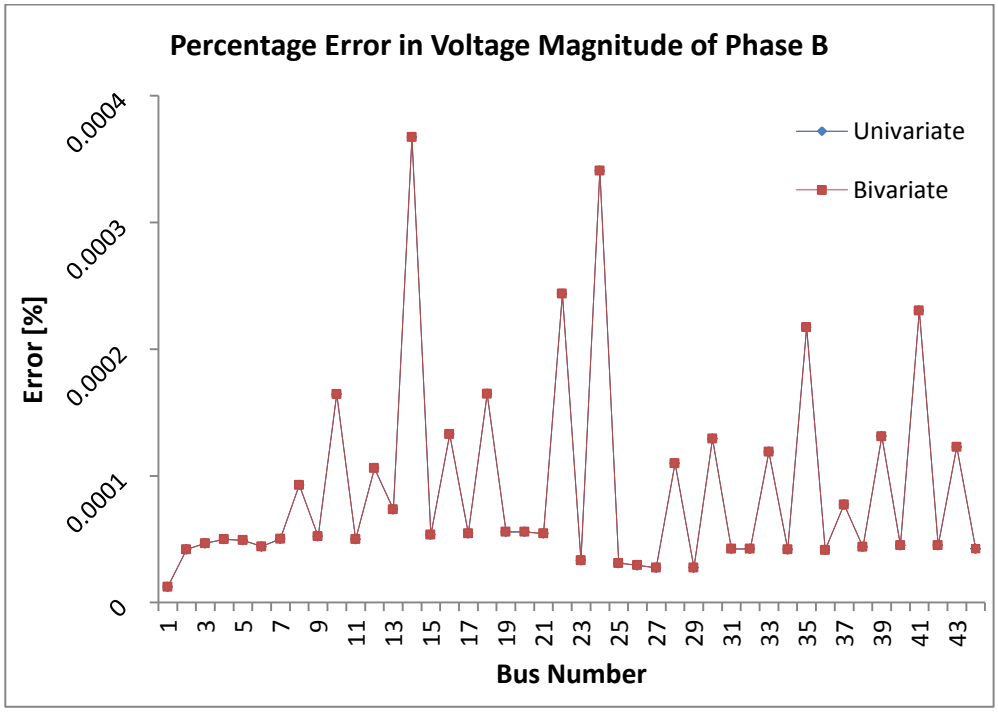

Fig 7.3: Average Percentage Error in Voltage Magnitude for Phase B Using the UT+UDR and UT+BDR Methods (*Both give almost the same results*)

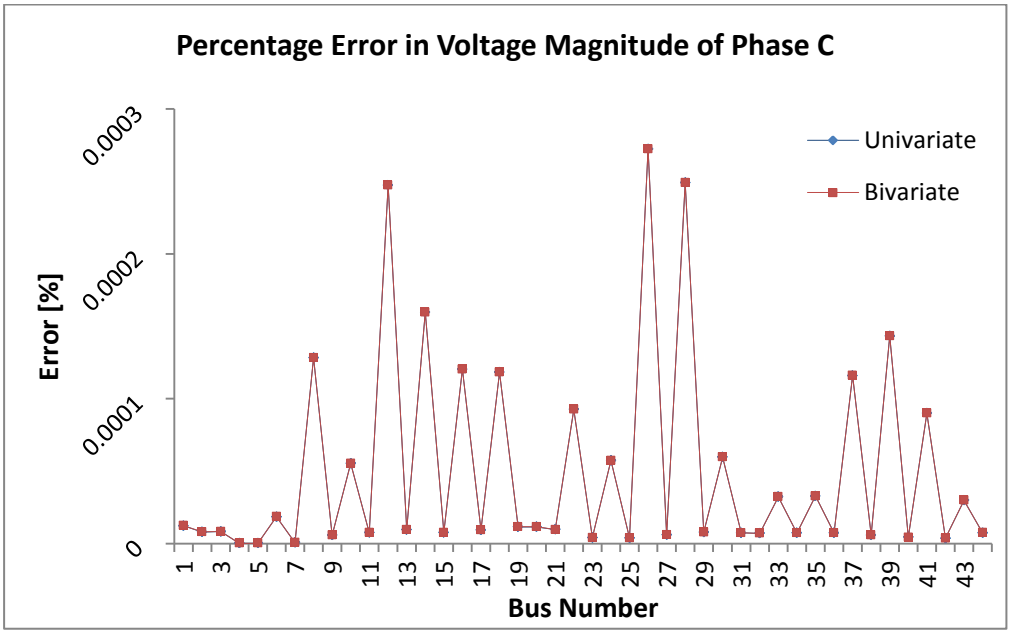

Fig 7. 4: Average Percentage Error in Voltage Magnitude for Phase B Using the UT+UDR and UT+BDR Methods (*Both give almost the same results*)

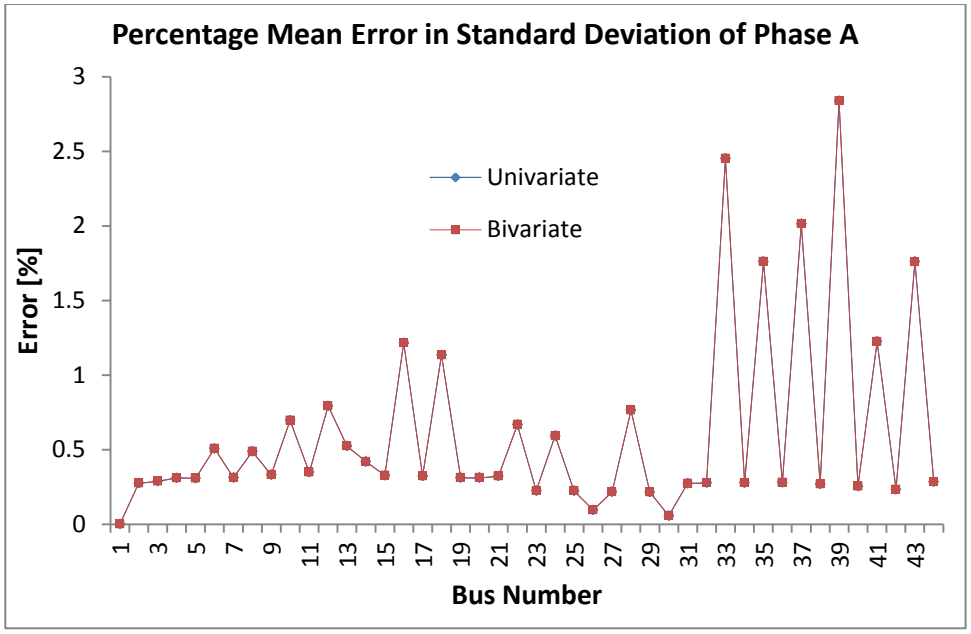

Fig 7.5: Average Percentage Error in the Standard Deviation of The Voltage Magnitude (Phase A) Using the UT+UDR and UT+BDR Methods (*Both give almost the same results*)

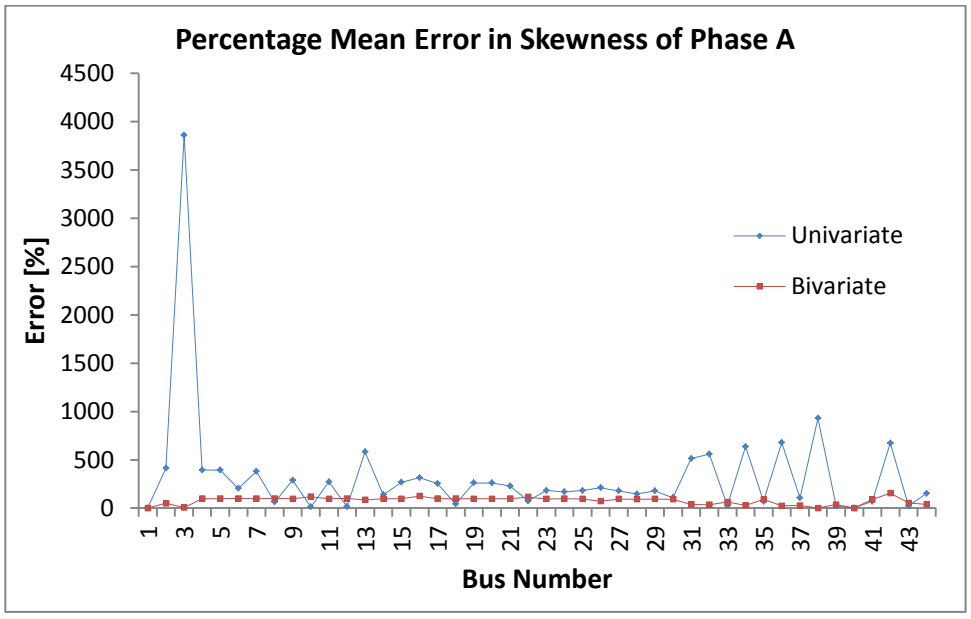

Fig 7.6: Average Percentage Error in the Skewness of The Voltage Magnitude (Phase A) Using the UT+UDR and UT+BDR Methods

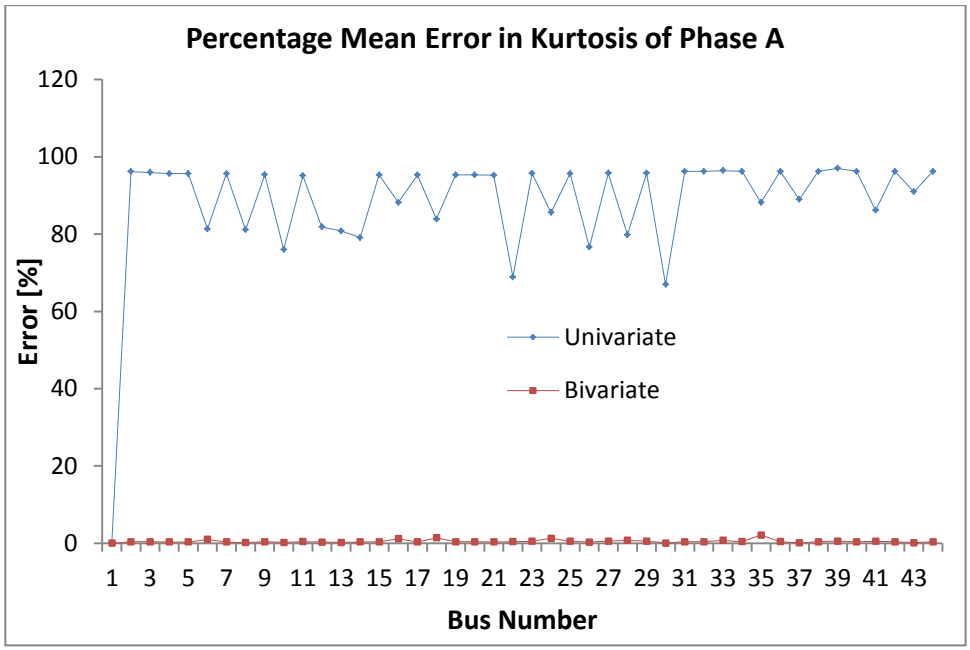

Fig 7.7: Average Percentage Error in the Kurtosis of The Voltage Magnitude (Phase A) Using the UT+UDR and UT+BDR Methods

From the graphs, it is seen that though the Univariate and bivariate dimension reduction based UT give approximately the same results for mean and standard deviation evaluation, the bivariate dimension reduction based UT provides a better estimate for higher order moments.

The voltage unbalance factor (VUF) is another important parameter for distribution systems. Voltage unbalance is defined as the ratio of the negative sequence component in the voltage to the positive sequence component in the voltage [7.7]. This is represented mathematically as (7.33);

$$VUF[\%] = \frac{Negative\ Sequence\ Voltage\ Component}{Positive\ Sequence\ Voltage\ Component} \times 100\% \qquad (7.33)$$

Following the EN50160 standard and other related standards like the EN61000-2-2, VUF should be maintained below 2% [7.8]. The percentage VUFs for all the buses using the Monte Carlo Simulation method, the univariate dimension reduction based UT and the bivariate dimension reduction based UT is plotted in Fig. 7.8. This figure shows the system conforms to the stipulated standard. However, it is seen that the proposed techniques give results similar to the MCS method with both the univariate and bivariate having about the same results. This amounts to a mean error of

1.2419×10^{-4} and 1.2418×10^{-4} for the Univariate and Bivariate dimension reduction based UT respectively, which is negligible. The errors here are within the range expected since the VUF evaluation involves only the mean of the voltage magnitude and angle with no higher order moments.

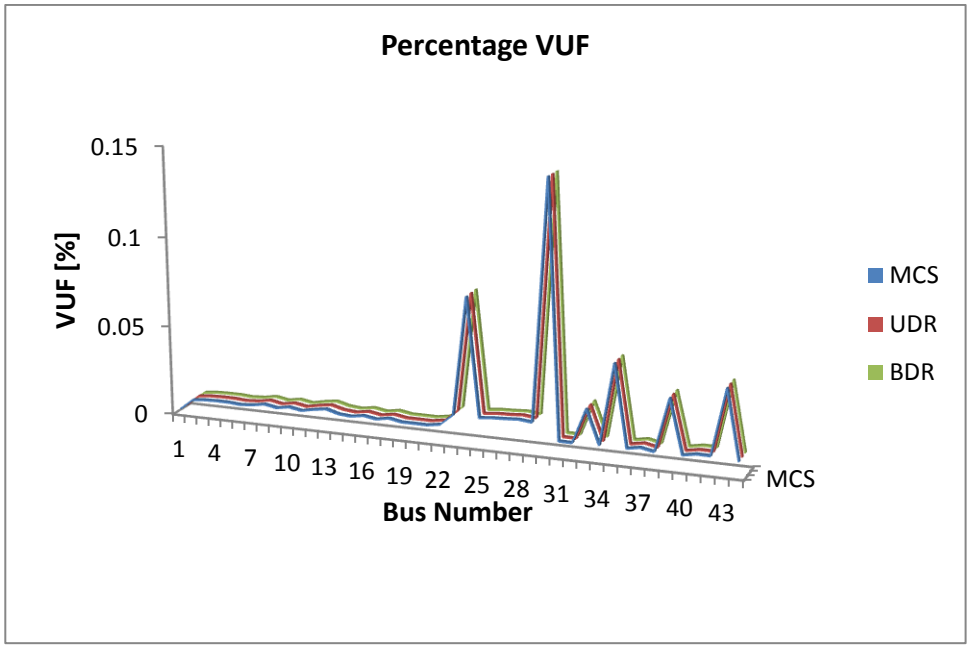

Fig 7.8: Percentage VUF using the Monte Carlo Simulation (MCS), Univariate Dimension Reduction based UT (UDR) and Bivariate Dimension Reduction Based UT (BDR) Methods

Note: EN50160 is the "voltage characteristics of electricity supplied by public distribution systems" standard while EN61000-2-2 is the "compatibility levels for low-frequency conducted disturbances and signalling in public low-voltage power supply systems" standard.

7.4. Précis

The dimension reduction technique has been introduced in this chapter to provide enhancement to the Gaussian quadrature based Unscented Transform method introduced in the previous chapter. Two variants of the technique namely the univariate and bivariate dimension reduction methods have been discussed including their mathematical basis. The performance of the techniques is compared with those for the Monte Carlo Simulation method, which is the established benchmark.

To evaluate the applicability of the techniques, three case studies involving two transmission systems and one real practical distribution system were studied. From the results, the limit of the Gaussian quadrature UT method in treating systems with large number of uncertainties despite its accuracy was reflected. It is seen that the bivariate dimension reduction based UT technique gives approximately the same result as with the Gaussian quadrature UT and Monte Carlo simulation method but requires less estimation points.

Both the univariate and bivariate dimension reduction based UT techniques give the same level of accuracy in estimating the mean and standard deviation of parameters, though, the Bivariate Dimension reduction based UT gives better estimate for higher order moments. This places it (bivariate dimension reduction) at a better position when higher order moments like the skewness and kurtosis are desired and in plotting the cumulative distribution function (and the probability density function). However, since the Bivariate Dimension Reduction based UT requires more points than its Univariate counterpart, more computational effort and time is required. In essence, the decision on which of the techniques to employ in a particular scenario depends on the most pressing quantity between higher order moments and speed/accuracy.

7.5. References

[7.1] H. Xu, and S. Rahman, "A generalized dimension-reduction method for multidimensional integration in stochastic mechanics," *Inter. Journal for Numerical Methods in Engineering*, pp. 1992-2019, Oct. 2004.

[7.2] I. Lee, K. K. Choi, L. Du, and D. Gorsich, "Dimension reduction method for reliability-based robust optimization," *Journal of Computers and Structures*, vol. 86, pp. 1550-1562, 2008.

[7.3] B. Huang, and X. Du, "Uncertainty Analysis by Dimension Reduction Integration and Saddlepoint Approximations," *Journal of Mechanical Design*, vol. 128, pp 26-33 Jan. 2006.

[7.4] A. Tamtum, A. Schellenberg, and D. Rosehart, "Enhancement to the Cumulant Method for Probabilistic Optimal Power Flow Studies," *IEEE Trans. Power Systems*, vol. 24, no 4, pp. 1739-1746, Nov. 2009.

[7.5] O. Oke, and D. W. P. Thomas, "Enhanced Cumulant Method for Probabilistic Power Flow in Systems with Wind Generation," In *proc. of International Conf. on Environment and Electrical Engineering EEEIC* 2012, pp. 849-853

[7.6] J.M. Morales, L. Baringo, A.J. Conejo and R. Minguez, " Probabilistic Power Flow with Correlated Wind Sources," *IET Generation Transmission and Distribution,* vol. 4, iss. 5, pp. 641-651, 2010

[7.7] P. Pillay, and M. Manyage, "Definitions of Voltage Unbalance," *IEEE Power Engineering Review,* May 2001, pp 50-51.

[7.8] *Voltage characteristics of electricity supplied by public distribution systems*, European Standard EN 50160, 2004

Chapter 8

Enhanced Unscented Transform and Dependency

The ideal alternative technique to the Monte Carlo Simulation method must be one able to take into account the effect of dependence which often exists amongst random variables under similar influences. To ensure that the proposed enhanced unscented transform meets the criterion, this chapter discusses a modification to the enhanced unscented transform to incorporate the effect of linear dependence (correlation). The mathematical basis for the principle is first discussed while its practicality is demonstrated in the final section for problems involving distributions following Gaussian and non-Gaussian distributions.

8.1. Power Systems and Dependency

Up till now, the effects of dependency among the uncertain variables have been neglected. However, considering that some form of dependency often exists between quantities in power systems their effects must be studied. For instance dependency may exist between the load demands of various customers in a particular region due to similar prevailing climatic and weather conditions or even similar work patterns. Also, there will be dependency between the outputs of wind farms located within the same area, since they will likely experience the same conditions (wind speed) due to proximity. Other form of dependency may also exist within the power system like those between the load demand and the output of generators. These effects must therefore be considered to get the true picture of the power system condition and also avoid wrong estimation and decisions [8.1].

Although dependence goes beyond linear dependence or correlation as already discussed in section (3.3.6) of chapter 3, only the effect of correlation is dealt with in this chapter since it is well established and has informed most load flow research studies focusing on this type of dependence

8.2. The Multivariate Dependent Problem and Possible Solutions

As discussed in earlier chapters, a problem involving more than one randomly varying parameter can be referred to as a multivariate problem. These are more common in real life and practical systems than univariate problems. So far, the methods used for the Unscented Transform method in Chapters 5 to 7 were based solely on the assumption of independency. To consider the effects of dependency, additional mathematical techniques are required. One possible solution is to incorporate the effect of linear dependency (correlation) using some form of transformation [8.2-8.13]. Another alternative will be to determine the orthogonal polynomial for the multivariate with the effect of correlation incorporated such that the zeroes of the polynomial become the sigma points just as discussed in Chapter 6 [8.3], [8.4]. The latter method is still under development and remains a research focus for mathematicians [8.5, 8.6]. The feasibility of employing the transformation technique is discussed in the following section.

8.3. Transformation

For a set of correlated variables, whose marginal distributions are known, the magnitude of correlation can be incorporated into their individual sigma points once the correlation coefficient (ρ) is given. The Cholesky decomposition provides a simple way of doing this. Another alternative is to use spectral decomposition [8.2]. The Cholesky decomposition is used in this work and its application to Univariate Dimension reduction based Unscented Transform is further discussed.

8.3.1. The Cholesky Decomposition for Correlation Incorporation

Using the quadrature method described in Chapter 6, equation (6.17) for two random variables can be written as (8.1) while the approximate equivalent is given as (8.2).

$$E\left[f(x_i, x_j)\right] = \int\int x_i x_j f(x_i, x_j) dx_i dx_j \tag{8.1}$$

$$E\left[f(x_i, x_j)\right] = \sum_{i=1, j=1}^{n,n} w_i w_j f(S_i, S_j) \tag{8.2}$$

Incorporating the effect of correlation using the correlation matrix *C*, (8.2) becomes [8.2];

$$E\left[f(x_i, x_j)\right] = \sum_{i=1, j=1}^{n,n} w_i w_j f(R_i, R_j) \tag{8.3}$$

where *R* is the transformed sigma points given by;

$$R = L.S \tag{8.4}$$

Based on Cholesky decomposition, *L* is the lower triangular matrix which satisfies (8.5).

$$C = LL^T \tag{8.5}$$

Applying the (8.3)-(8.5) above gives a simple solution for the two variable case.

$$E\left[f(x_i, x_j)\right] = \sum_{i=1, j=1}^{n,n} w_i w_j f(S_i, \rho S_i + \rho' S_j) \tag{8.6}$$

$$\rho' = \sqrt{1 - \rho^2} \tag{8.7}$$

Since correlation does not affect the mean of a variable, the mean of each correlated variable is deducted before the effect of correlation is included.

The above method can be simply extended to the Univariate Dimension Reduction (UDR) based UT. Using the UDR expansion from (7.9) rewritten here as (8.8), the first part of the equation which is a function of the variable ***x*** is multiplied by the lower triangular matrix *L*.

$$\hat{f}(\mathbf{x}) = \sum_{i=1}^{N} f(\bar{x}_1, \bar{x}_2, \ldots, x_i, \bar{x}_N) - (N-1) f(\bar{\mathbf{x}}) \tag{8.8}$$

As an illustration, for a problem involving three correlated random variables whose marginal distributions are known, the decomposition following the UDR method is given as;

$$\begin{aligned}\hat{f}(\mathbf{R}) = \hat{f}(L\cdot\mathbf{x}) = f(L\cdot(x_1,\bar{x}_2,\bar{x}_3)) + f(L\cdot(\bar{x}_1,x_2,\bar{x}_3))\ldots \\ + f(L\cdot(\bar{x}_1,\bar{x}_2,x_3)) - 2f(\bar{x}_1,\bar{x}_2,\bar{x}_3)\end{aligned} \tag{8.9}$$

For ease, the first part of the equation which is a function of **x** can be written in matrix form as;

$$L\cdot(x_1,\bar{x}_2,\bar{x}_3) = L\cdot\begin{bmatrix} x_1 & \bar{x}_2 & \bar{x}_3 \\ \bar{x}_1 & x_2 & \bar{x}_3 \\ \bar{x}_1 & \bar{x}_2 & x_3 \end{bmatrix} \tag{8.10}$$

Since the mean of a variable is not affected by correlation, the means can be deducted before accounting for the correlation. Matrix (8.10) then becomes;

$$L\cdot(\hat{x}) = L\cdot\begin{bmatrix} \hat{x}_1 & 0 & 0 \\ 0 & \hat{x}_2 & 0 \\ 0 & 0 & \hat{x}_3 \end{bmatrix} \tag{8.11}$$

where $\hat{x} = x - \bar{x}$

Let the lower triangular matrix of the Cholesky decomposition be represented by (8.12).

$$L = \begin{bmatrix} 1 & 0 & 0 \\ a & b & 0 \\ c & d & e \end{bmatrix} \tag{8.12}$$

Then the transformed correlated form (L·**x**) becomes;

$$L\cdot\hat{\mathbf{x}} = \begin{bmatrix} \hat{x}_1 & 0 & 0 \\ a\hat{x}_1 & b\hat{x}_2 & 0 \\ c\hat{x}_1 & d\hat{x}_2 & e\hat{x}_3 \end{bmatrix} \tag{8.13}$$

(8.13) is then substituted back into (8.9) with the addition of the mean.

The Cholesky decomposition technique can also be applied to Monte Carlo simulation involving Gaussian distributed variables. In problems involving non-Gaussian distribution, additional techniques are required for the MCS method in order to preserve the original distribution of the variables while incorporating correlation effects. Alternative technique using Nataf transformation is discussed below.

8.3.2. The Nataf Transformation

In situations where the marginal distributions and the covariance matrix are available, the Nataf transformation technique [8.3] can easily be employed to account for the effect of correlation.

For *n* dependent random variables $\mathbf{X}=(X_1,X_2,\ldots,X_n)$ with known marginal distributions and correlation matrix $\boldsymbol{\rho_0}=\{\rho_{ij}\}$, the marginal transformation to standard normal variables $Y=(Y_1,Y_2,\ldots,Y_n)$ in the *y* space is given by;

$$Y_i = \Phi^{-1}[F_{Xi}(xi)] \quad i = 1,\ldots,n \tag{8.14}$$

where Φ^{-1} is the inverse cumulative distribution function of a standard normal variable.

Following the Nataf transformation, the approximate joint density function $f_X(X_1,X_2,\ldots,X_n)$ in the **x** space is given by (8.15) [8.3].

$$f_X(\mathbf{X}) = f_{X_1}(x_1) f_{X_2}(x_2)\cdots f_{X\,n}(x_n)\frac{\phi_n(y,\rho_0)}{\phi(y_1)\phi(y_2)\ldots\phi(y_n)} \tag{8.15}$$

where

$$\phi_n(y,\rho_0) = \frac{1}{\sqrt{(2\pi)^n \det(\rho_0)}}\exp(-\frac{1}{2}y^T\rho_0 y) \tag{8.16}$$

Equation (8.16) is the referred to as the Nataf transformation. To solve for the components of the correlation matrix, (8.17) is employed [8.3].

$$\rho_{ij} = \frac{\text{cov}[X_i X_j]}{\sigma_{X_i}\sigma_{X_j}} = E[Z_i Z_j] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{F_{X_i}^{-1}(\Phi(y_i)) - \mu_{X_i}}{\sigma_{X_i}}\right)\cdots \left(\frac{F_{X_j}^{-1}(\Phi(y_j)) - \mu_{Xj}}{\sigma_{X_{ji}}}\right)(y_i, y_j, \rho_{oij})dy_i dy_j \qquad (8.17)$$

where ρ_{0ij} is a component in the correlation matrix of the standard normal random vector **Y** which is related to the correlation matrix by (8.18).

$$R = \rho_{oij} / \rho_o \qquad (8.18)$$

$$R = a + bV_i + cV_i^2 + d\rho + e\rho^2 + f\rho V_i + gV_j + hV_j^2 + k\rho V_j + lV_i V_j \qquad (8.19)$$

The values of the various components of R for several distributions are presented in [8.3]. V_i and V_j are the coefficient of variation (CV) of the variable i and j respectively (as defined in Section 3.3.3, Chapter 3).

This transformed correlation matrix is employed for non-Gaussian distributed random variables.

8.5. Implementation Procedure

The procedure for applying the Enhanced Unscented Transform method which includes all the proposed enhancements are summarized below while the flowchart is presented in Fig 8.1.

i. Identify all randomly varying distributions within the power system. Characterise and model the variables (based on past data) to obtain their probability density function (PDF). [Chapter 3]

ii. Compute the initial sigma points and weights for all the random variables using Gaussian Quadrature, Orthogonal polynomial, Stieltjes procedure and the triadiagonal Jacobian matrix. [Chapter 6]

iii. For multivariate problems, apply dimension reduction technique (Univariate or Bivariate) in combining the sigma points and weights obtained above. [Chapter 7]

iv. Get the correlation matrix to incorporate the effect of correlation. Decompose the matrix using the Cholesky decomposition technique. Transform the sigma points above using the lower triangular matrix from the decomposition. [Chapter 8]

v. Run load flow using the final sigma points and weights. Obtain the moments and statistical data required. [Chapter 3]

vi. Plot the required output cumulative distribution function and probability density function using the Cornish Fisher Method, Gram Charlier or Pearson Curve. [Chapter 3].

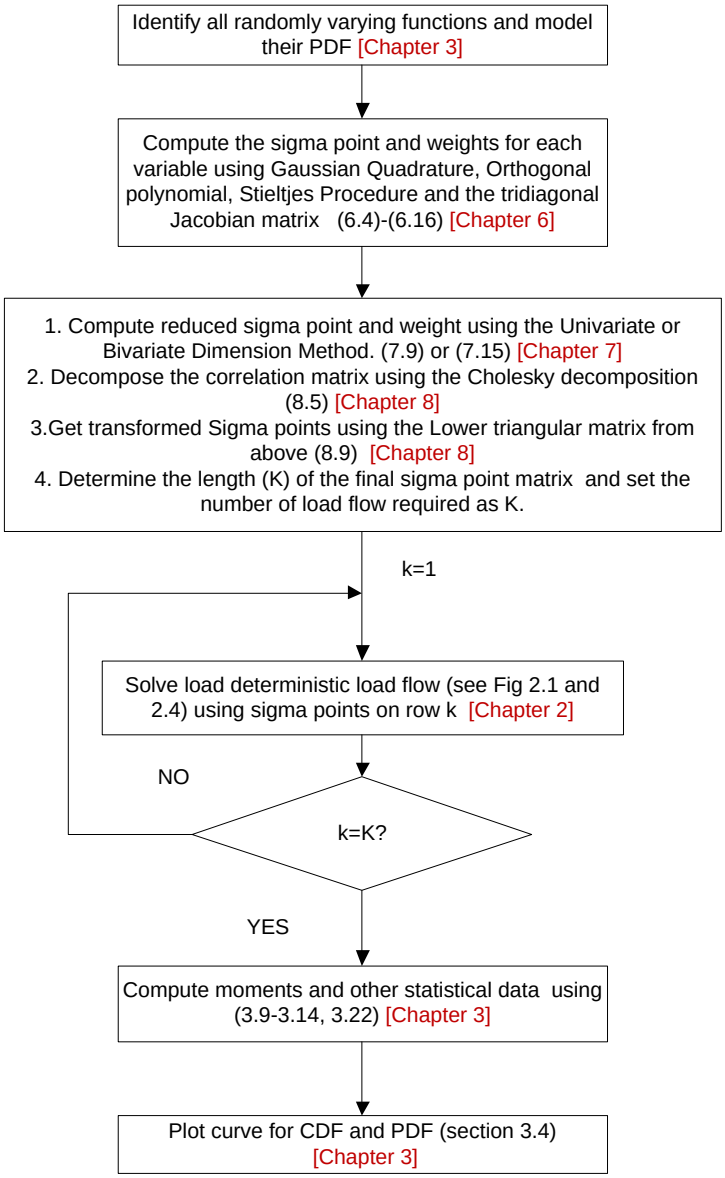

Fig 8.1: Flowchart for the Implementation of the Enhanced Unscented Transforms Method

8.6. Case Studies

In this section, the performance of the methods described above is evaluated using systems with both Gaussian and non-Gaussian random variables. In all, three case studies are examined. The first is aimed at understanding the effect of correlation on the power system while the second examines the impact (if any), of power system size on the accuracy of the proposed technique. Considering that the first two case studies are based on having Gaussian correlated variables only, the last case study evaluates the performance of the technique for non-Gaussian distributed variables which is typical of the wind speed and power distributions.

In all the cases, the performance of the proposed method is compared against those from the Monte Carlo Simulation method and the $2n+1$ point estimate method.

8.6.1. Case 8.1: The Effect of Accounting for Correlation in Load Flow Studies

In this section, load flow studies are carried out for a simple system with and without considering correlation. A simple 6-bus system (Appendix C) is employed in this case. The active power on buses 5 and 6 with mean 70MW and 55MW respectively have 5% coefficient of variation. For the second case involving correlated variables, the two random variables are assumed to have a correlation coefficient (ρ) of 0.8.

For the scenario involving correlation, the technique on Cholesky decomposition discussed above is employed for the univariate dimension reduction based UT (UT+UDR) and the Monte Carlo Simulation (MCS) methods while the technique proposed in [8.13] are used for the point estimate method (PEM). In both cases, 10,000 samples are used for the Monte Carlo Simulation while 3 sigma points are used in representing each variable for the UT+UDR method. The UT+UDR method requires 7 estimations while the PEM uses 5 estimation points in all. It is worth stating that the reduced UT+UDR using three points require the same number of evaluation as the

PEM and gives exactly the same results as the UT+UDR., but it can only be used for symmetrical distributions with one sigma point located on the mean.

The results obtained for both scenarios using the three methods are presented in Tables 8.1-8.8.

Table 8.1: Selected Results for Voltage Magnitude (Without Correlation)

Moments (Voltage)		MCS	UT+UDR	PEM
V_4	μ	0.98893	0.98893	0.98893
	σ	9.27e-05	9.29e-05	9.29e-05
V_5	μ	0.985985	0.98598	0.98598
	σ	9.67e-4	9.68e-4	9.68e-4
V_6	μ	1.00725	1.00725	1.00725
	σ	5.43e-4	5.44e-4	5.44e-4

Table 8.2: Selected Results for Voltage Magnitude (With Correlation)

Moments (Voltage)		MCS	UT+UDR	PEM
V_4	μ	0.98893	0.98893	0.98893
	σ	6.09e-05	7.27e-05	7.56e-05
V_5	μ	0.98598	0.98598	0.98598
	σ	1.035e-3	1.036e-3	9.66e-4
V_6	μ	1.00725	1.00725	1.00725
	σ	6.749e-4	7.76e-4	4.40e-3

From the results, it is seen that including correlation gives a different results as compared to when it was ignored considering all quantities examined (voltage magnitude and angle, active and reactive power). This effect is more glaring for the standard deviation while the mean (expectation) remains fairly constant. This corroborates the initial statement that correlation does not impact on the

mean of a variable. Therefore, if the true system state is desired, the effect of correlation must be fully taken into account while carrying out load flow studies.

Table 8.3: Selected Results for Voltage Angle (Without Correlation)

Moments (Voltage Angle)		MCS	UT+UDR	PEM
δ_4	μ	-3.63525	-3.63505	-3.63505
	σ	0.156134	0.1563	0.1563
δ_5	μ	-4.44125	-4.44109	-4.44109
	σ	0.29913	0.29941	0.29941
δ_6	μ	-4.45055	-4.44981	-4.44981
	σ	0.33747	0.33794	0.33794

Table 8.4: Selected Results for Voltage Angle (With Correlation)

Moments (Voltage Angle)		MCS	UT+UDR	PEM
δ_4	μ	-3.6359	-3.6353	-3.635
	σ	0.20887	0.23060	0.14238
δ_5	μ	-4.44261	-4.44142	-4.44102
	σ	0.38989	0.42284	0.28354
δ_6	μ	-4.45153	-4.45028	-4.44969
	σ	0.44779	0.50448	0.29061

Table 8.5: Selected Results for Active Power Flow (Without Correlation)

Moments (Active Power)		MCS	UT+UDR	PEM
$P_{4,1}$	μ	0.37841	0.37840	0.37840
	σ	0.01306	0.01307	0.01307
$P_{6,2}$	μ	0.19638	0.19634	0.19634
	σ	0.00785	0.00786	0.00786
$P_{6,3}$	μ	0.37467	0.37463	0.37463
	σ	0.00971	0.00974	0.00974

Table 8.6: Selected Results for Active Power Flow (With Correlation)

Moments (Active Power)		MCS	UT+UDR	PEM
$P_{4,1}$	μ	0.37846	0.37841	0.37839
	σ	0.01746	0.01927	0.01191
$P_{6,2}$	μ	0.19638	0.19634	0.19634
	σ	0.01032	0.01676	0.00962
$P_{6,3}$	μ	0.37466	0.37463	0.37463
	σ	0.00929	0.01104	0.00756

Table 8.7: Selected Results for Reactive Power Flow (Without Correlation)

Moments (Reactive Power)		MCS	UT+UDR	PEM
$Q_{4,1}$	μ	0.21645	0.21646	0.21646
	σ	0.00431	0.00431	0.00431
$Q_{6,2}$	μ	0.16992	0.16993	0.16993
	σ	0.00191	0.00191	0.00191
$Q_{6,3}$	μ	0.56406	0.56405	0.56405
	σ	0.00328	0.00328	0.00328

Table 8.8: Selected Results for Reactive Power Flow (With Correlation)

Moments (Reactive Power)		MCS	UT+UDR	PEM
$Q_{4,1}$	μ	0.21643	0.21644	0.21646
	σ	0.00578	0.00643	0.00384
$Q_{6,2}$	μ	0.16992	0.16992	0.16993
	σ	0.00237	0.00273	0.00155
$Q_{6,3}$	μ	0.56407	0.56406	0.56405
	σ	0.00436	0.00491	0.00284

8.5.2. Case 8.2: Performance Analysis for variables Following the Gaussian Distribution

This section examines the performance of the UT+UDR for systems with correlated variables. To ensure the general applicability of the technique, its performance on 3 test systems is evaluated. The previously employed 6-bus

test system and the IEEE 14-bus and 118-bus test systems are studied. The 6-bus test system is assumed to have the same parameters as in case 1 above.

The standard IEEE 14-bus test system (Appendix C) is modified such that some buses similar to those in section 7.3.2 (expect for the exclusion of the wind farm and the generator forced outage rates) are made to vary randomly. The system includes 22 random variables all with a 5% coefficient of variation. In the simulation involving correlated random variables, two areas have been identified; the first includes the active power on buses 2, 3 and 4 while area two comprises of buses 5, 6 and 9. The three buses in the same area are assumed to have a correlation coefficient of 0.6 while those in different areas have a correlation coefficient of 0.2. The correlation matrix is shown below. For the MCS method, 10,000 sample points are used while 67 points were used for the UT+UDR (3 points). The PEM employed 45 points as shown in Table 8.9.

$$[\rho] = \begin{bmatrix} 1 & 0.6 & 0.6 & 0.2 & 0.2 & 0.2 \\ 0.6 & 1 & 0.6 & 0.2 & 0.2 & 0.2 \\ 0.6 & 0.6 & 1 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 1 & 0.6 & 0.6 \\ 0.2 & 0.2 & 0.2 & 0.6 & 1 & 0.6 \\ 0.2 & 0.2 & 0.2 & 0.6 & 0.6 & 1 \end{bmatrix}$$

The IEEE 118-test system (Appendix C) is modified to include 82 random variables all following the Gaussian distribution with a 5% coefficient of variation. As with the 14-bus test system above, two areas have been identified for the IEEE 118-bus test system. The first comprises of buses 74, 76 and 77 and the second area includes buses 80, 85 and 90. The correlation is assumed to be exactly as with the 14-bus system.

Table 8.9: Number of Simulation for Each System and Technique

Test System	MCS	UT+UDR (3pts)	PEM
6 Bus	10000	7	5
14-Bus	10000	67	45
118-Bus	10000	246	165

N.B. The reduced UT+UDR can also be employed here since the problem involves only Gaussian distribution. The results will be the same as the UT+UDR but while the number of evaluation required will be the same of those for PEM.

The average percentage error in the mean and standard deviation for the 6-bus test system without considering correlation is presented in Table 8.10 while Table 8.11 shows the errors in the system when the correlation between the variables are considered. UT+UDR and PEM have been compared against the MCS benchmark.

Table 8.10: Average Percentage Error in Mean and Standard Deviation for the 6-Bus Test System (Without Correlation)

Quantity/Parameter	Moment	UT+UDR	PEM
Voltage Magnitude	ε_μ	5.73e-5	5.73e-5
	ε_σ	0.16524	0.16524
Voltage Angle	ε_μ	0.01084	0.01084
	ε_σ	0.11418	0.11418
Active Power Flow	ε_μ	0.06839	0.06839
	ε_σ	0.15301	0.15301
Reactive Power Flow	ε_μ	0.00397	0.00397
	ε_σ	0.11344	0.11344

Table 8.11: Average Percentage Error in Mean and Standard Deviation for the 6-Bus Test System (With Correlation)

Quantity/Parameter	Moment	UT+UDR	PEM
Voltage Magnitude	ε_μ	6.87e-5	2.07e-4
	ε_σ	12.1267	21.8607
Voltage Angle	ε_μ	0.02252	0.04008
	ε_σ	9.02555	32.2290
Active Power Flow	ε_μ	0.10318	0.12109
	ε_σ	11.1203	35.4717
Reactive Power Flow	ε_μ	0.00614	0.02172
	ε_σ	11.1726	31.6627

From Table 8.10, the UT+UDR and PEM give the same level of average percentage errors for the mean and standard deviation. This is expected since all the variables here follow the Gaussian distribution. Generally, it is seen that the error in the mean estimation is much lower than those in the standard

deviation for both methods, though, the magnitude of the errors for the correlated case is higher. For the system with correlated variables, the UT+UDR give a better performance than the PEM for both the mean and standard deviation.

To further evaluate the performance of the UT+UDR, the results obtained when the method is applied to the 14-bus test system are presented in Tables 8.12 and 8.13.

Table 8.12: Average Percentage Error in Mean and Standard Deviation for the 14-Bus Test System (Without Correlation)

Quantity/Parameter	Moment	UT+UDR	PEM
Voltage Magnitude	ε_μ	5.64e-4	5.64e-4
	ε_σ	1.44763	1.45205
Voltage Angle	ε_μ	0.02611	0.02611
	ε_σ	0.77463	0.78591
Active Power Flow	ε_μ	0.05390	0.05390
	ε_σ	0.63674	0.63652
Reactive Power Flow	ε_μ	0.01953	0.01953
	ε_σ	5.98012	5.96330

Table 8.13: Average Percentage Error in Mean and Standard Deviation for the 14-Bus Test System (With Correlation)

Quantity/Parameter	Moment	UT+UDR	PEM
Voltage Magnitude	ε_μ	1.33e-4	4.55e-4
	ε_σ	2.20065	12.1460
Voltage Angle	ε_μ	0.04093	0.04598
	ε_σ	2.43862	3.82323
Active Power Flow	ε_μ	0.04570	0.04714
	ε_σ	8.23629	17.6135
Reactive Power Flow	ε_μ	0.12153	0.12047
	ε_σ	8.90074	16.1760

As with the 6-bus test system, the results for the 14-bus test system without considering the effects of correlation show that the PEM and UT+UDR give about the same level of accuracy. For the scenario with the correlated variables,

the UT+UDR gives better accuracy than the PEM. The magnitude of the errors in the standard deviation of the various parameters is lower as compared with the 6-bus test system, although this may be because all variables within the latter are correlated (full correlation) why about 30% of the variables in the 14-bus system are correlated.

To further prove the applicability of the proposed method, the performance of the method in a larger realistic IEEE 118-bus test system is presented in Tables 8.14 and 8.15 in terms of the percentage errors.

Table 8.14: Average Percentage Error in Mean and Standard Deviation for the 118-Bus Test System (Without Correlation)

Quantity/Parameter	Moment	UT+UDR	PEM
Voltage Magnitude	ε_μ	1.34e-5	1.35e-5
	ε_σ	0.30676	0.30675
Voltage Angle	ε_μ	0.01046	0.01046
	ε_σ	0.52313	0.52314
Active Power Flow	ε_μ	0.06142	0.06142
	ε_σ	0.39166	0.39166
Reactive Power Flow	ε_μ	0.02951	0.02951
	ε_σ	0.95754	0.95754

Table 8.15: Average Percentage Error in Mean and Standard Deviation for the 118-Bus Test System (With Correlation)

Quantity/Parameter	Moment	UT+UDR	PEM
Voltage Magnitude	ε_μ	3.84e-5	2.47e-5
	ε_σ	1.23602	1.82245
Voltage Angle	ε_μ	0.02337	0.02300
	ε_σ	1.53370	4.72917
Active Power Flow	ε_μ	0.06732	0.06812
	ε_σ	1.90991	2.44122
Reactive Power Flow	ε_μ	0.04693	0.04665
	ε_σ	2.48716	3.40042

From the results, as with the previous test systems, both UT+UDR and PEM give close results for the uncorrelated case while the former performs better in

estimating the standard deviation of the correlated case. It is seen that the UT+UDR method gives a good approximation to the MCS method since the maximum average errors in both the correlated and the uncorrelated scenarios is less than 3% while that of the PEM is about 5%.

8.6.3. Case 8.3: Accounting for Correlation In Systems with Wind Generation

To ensure the applicability and adaptability of the UT+UDR method for load flow related problems, its performance must be examined for problems involving variables with both Gaussian and non-Gaussian distributions which is typical of practical systems. In this section, the performance of the method is evaluated for a system involving two wind farms in close proximity whose wind speed and hence output power are correlated. The wind farms are assumed to be located on buses 55 and 90 with rated power of 50MW and 170MW respectively (the wind speed and wind turbine parameters are assumed to be the same as those employed in previous chapters).

As in case 8.2, initially, the performance of the system with 82 randomly distributed parameters (80 variable load and 2 wind farms) is examined to understand the performance of the system without considering correlation. After this, the outputs of the two wind farms are assumed to be correlated with coefficient of correlation being 0.9. Selected results for cases with large errors from the respective scenarios are presented in Tables 8.16 and 8.17.

Table 8.16: Selected Results for the 118-Bus Test System (Without Correlation)

Moments		MCS	UDR	PEM
Mean	V_{53}	0.94653	0.94653	0.94653
	V_{97}	1.00783	1.00784	1.00795
	$P_{68,69}$	0.10051	0.10335	0.12540
	$P_{68,81}$	1.1900	1.18773	1.17244
Std	V_{53}	0.00024	0.00023	0.00022
	V_{97}	0.00224	0.00226	0.00218
	$P_{68,69}$	0.28785	0.23510	0.22203
	$P_{68,81}$	0.18341	0.20317	0.19205

Table 8.17: Selected Results for the 118-Bus Test System (With Correlation)

Moments		MCS	UDR	PEM
Mean	V_{53}	0.94654	0.94653	0.94652
	V_{97}	1.00779	1.00787	1.00829
	$P_{68,69}$	0.98033	0.10308	0.16536
	$P_{68,81}$	1.19164	1.18803	1.12599
Std	V_{53}	0.00024	0.00023	0.00022
	V_{97}	0.00232	0.00195	0.00023
	$P_{68,69}$	0.2790435	0.22499	0.16359
	$P_{68,81}$	0.18866	0.17459	0.06170

For a thorough evaluation of the performance of the proposed method to be carried out, the percentage errors in the mean and standard deviation of each bus (voltage magnitude, voltage angle) and line (active power and reactive power flow) in both scenarios are presented in Fig. 8.2-8.5. The MCS is used as the basis for comparison while the errors in the PEM are presented alongside.

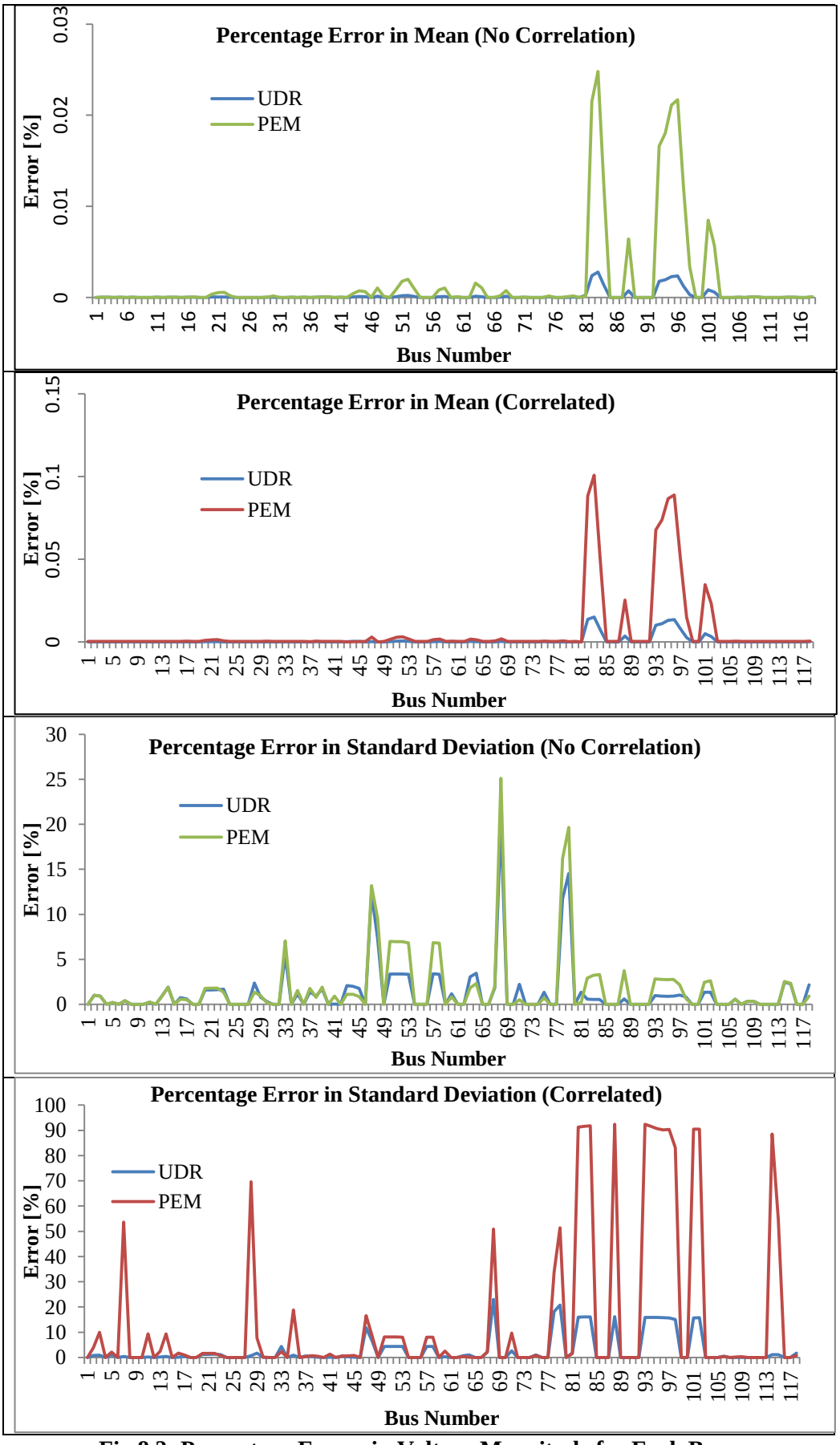

Fig 8.2: Percentage Errors in Voltage Magnitude for Each Bus

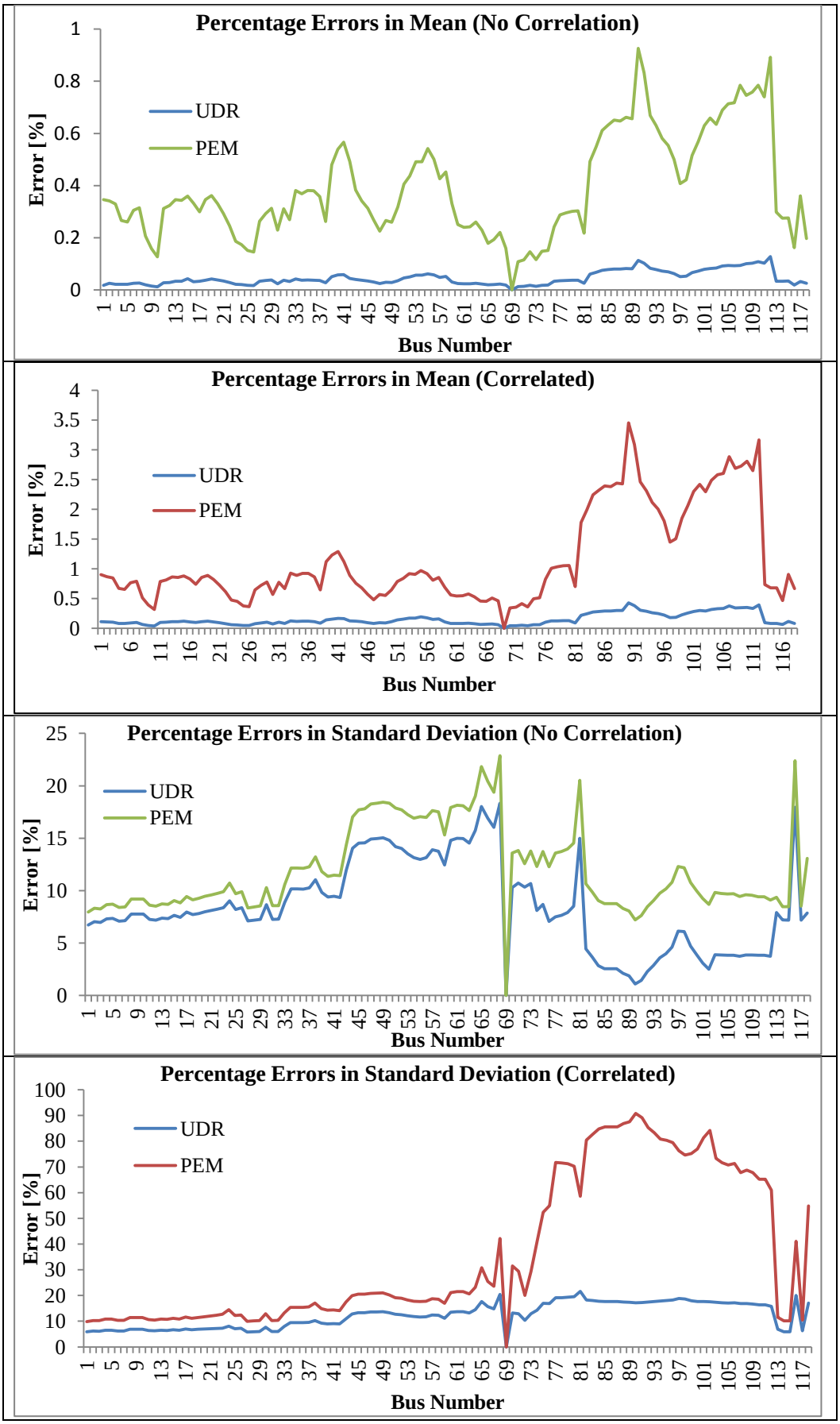

Fig 8.3: Percentage Errors in Voltage Angle for Each Bus

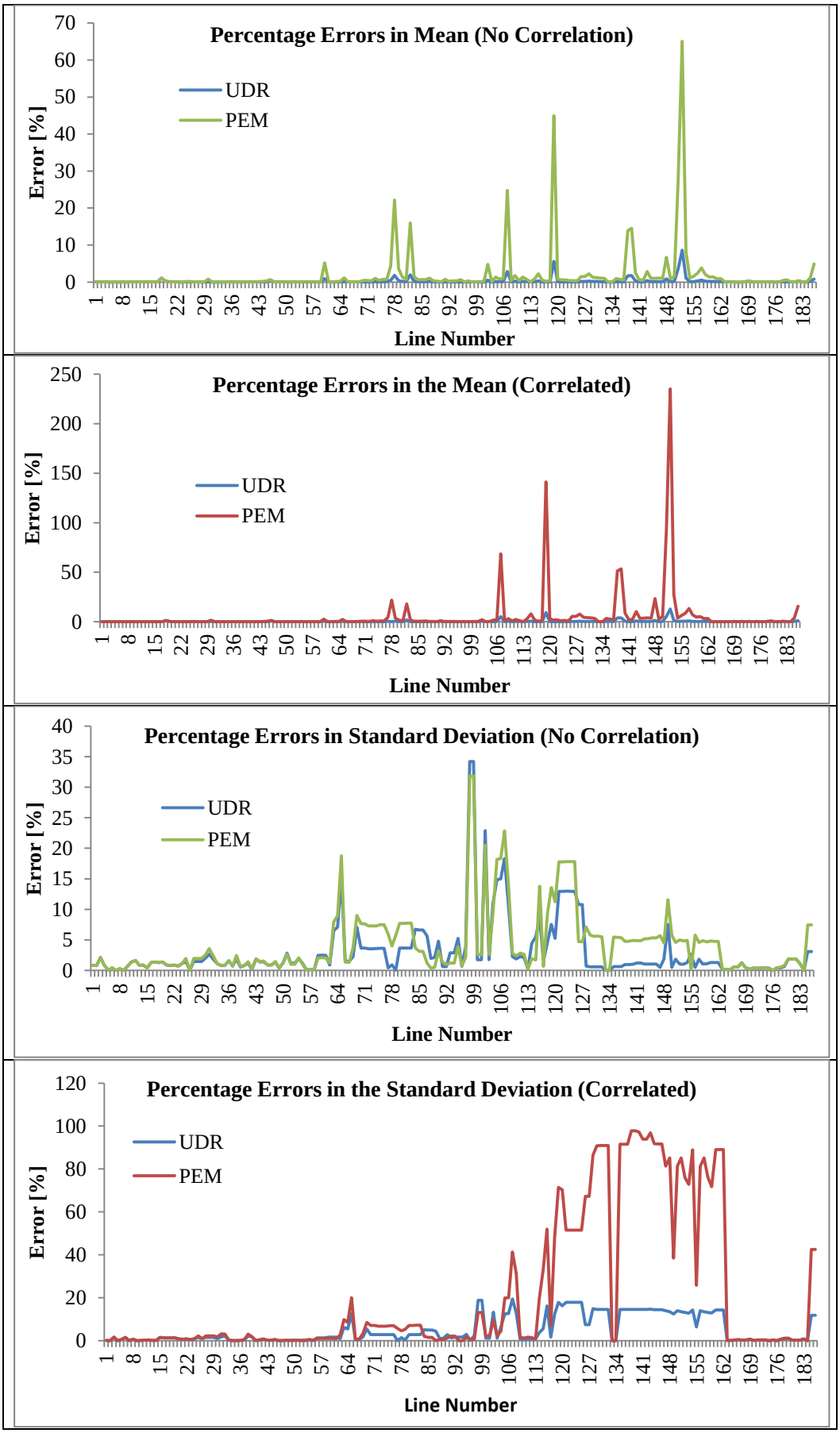

Fig 8.4: Percentage Errors in Active Power Flow for Each Line

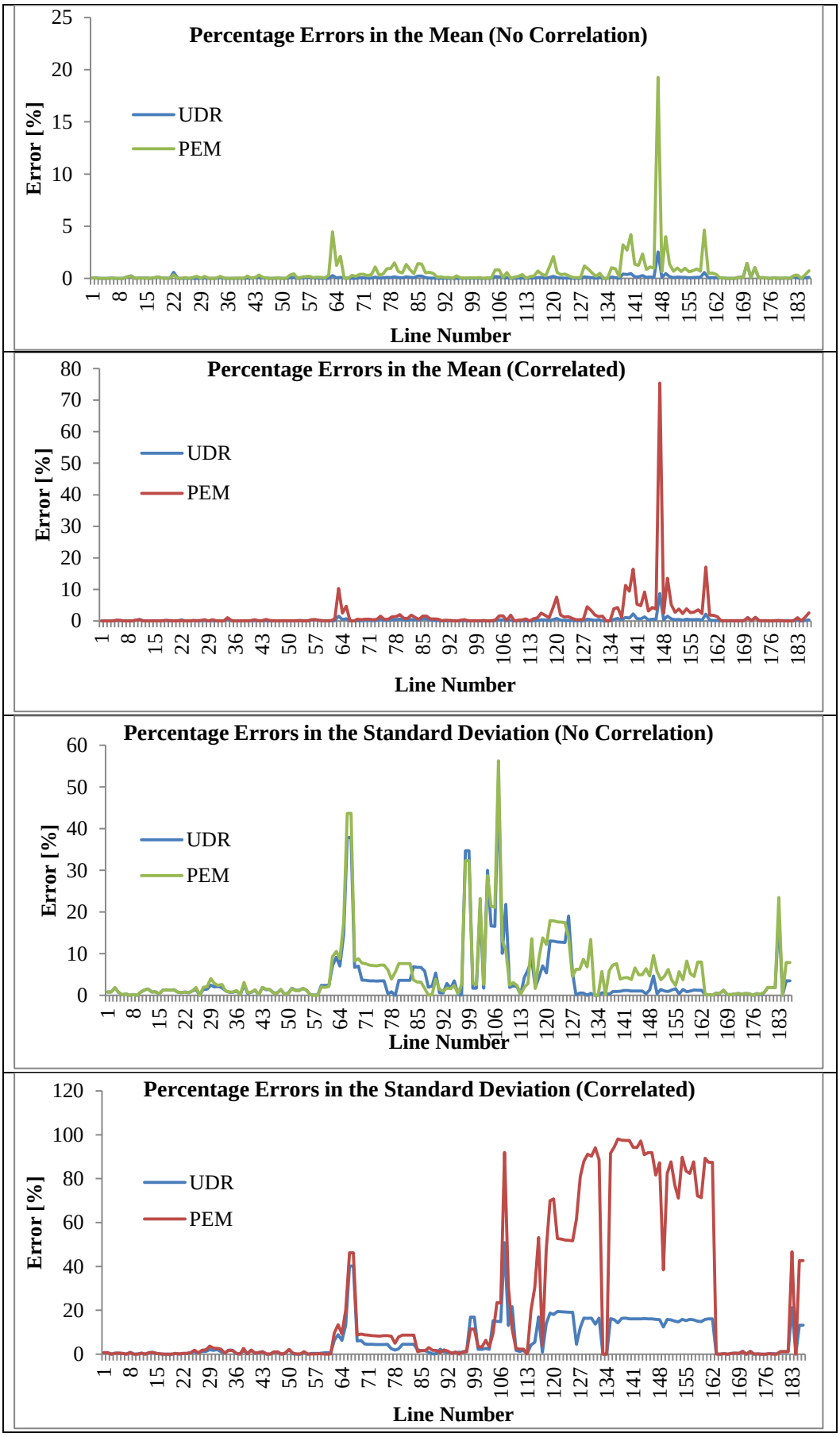

Fig 8.5: Percentage Errors in Reactive Power Flow for Each Line

For ease of explanation the scenario where correlation between the wind farms is neglected will be referred to as Scenario A while the other scenario where it is incorporated will be called Scenario B. From Fig. 8.2, it is seen that for the Scenario A, the errors are generally lower as compared with the Scenario B. A similar trend is also noticed in Fig. 8.3-8.5. The overall average percentage errors are shown in Tables 8.18 and 8.19 while the computation time using all the three methods is shown in Table 8.20.

The errors in the mean of the voltage magnitudes are similar or lower using the proposed UT+UDR method than the PEM for all the buses in both scenarios. For instance, the maximum error obtained for the UT+UDR method for standard deviation evaluation was ~20% and ~23% for Scenarios A and B respectively while the for the PEM the errors were ~25% and ~92%. Although, considering that the variations in the voltage magnitude are generally very low, these errors are very trivial as the largest relative error was 1.66E-5. In terms of the computation time, the PEM is about 57 times faster than the MCS in this particular case while the UT+UDR is about 38 times faster than the MCS. The PEM is seen to be about 1.5 times faster than the proposed UT+UDR, however, the UT+UDR can be as much as 3 times more accurate than the PEM, especially for the correlated scenario.

Overall, the UT+UDR method remains a better alternative with its ability to give good level of accuracy within a reasonable time even for evaluating parameters involving very small numbers.

Table 8.18: Overall Average Percentage Errors for Scenario A

Parameter		UT+UDR	PEM
Voltage Magnitude	ε_μ	1.76e-4	1.60e-3
	ε_σ	1.29643	1.74358
Voltage Angle	ε_μ	0.04481	0.38785
	ε_σ	8.45583	11.9482
Active Power Flow	ε_μ	0.24983	1.86594
	ε_σ	3.01714	4.22219
Reactive Power Flow	ε_μ	0.09492	0.56804
	ε_σ	3.89802	5.41123

Table 8. 19: Overall Average Percentage Errors for Scenario B

Parameter		UT+UDR	PEM
Voltage Magnitude	ε_μ	9.61e-4	6.45e-3
	ε_σ	2.79868	14.0764
Voltage Angle	ε_μ	0.15063	1.14945
	ε_σ	12.3171	35.7773
Active Power Flow	ε_μ	0.36868	5.24413
	ε_σ	5.04398	21.2921
Reactive Power Flow	ε_μ	0.26140	1.70559
	ε_σ	6.35523	23.0376

Table 8.20: Simulation Time Using The MCS, UT+UDR and PEM

MCS (sec)	UT+UDR (sec)	PEM (sec)
890.30	23.608	15.561

Considering Fig 8.3 for the voltage angle, a similar trend to those for the voltage magnitude is noticed. The magnitude of the errors here cannot be downplayed. For instance, for Scenario B, the maximum percentage error in the standard deviation for the estimation using the UT+UDR and PEM are 22% and 91% which amounts to 0.9^o and 4.7^o respectively. This large variation using the PEM can lead to error in both the design and operation stages of the power system.

This error becomes even more significant when considering the power flow in the lines, with the maximum error for Scenario B being ~13% and ~235% for the UT+UDR and PEM respectively. Since this error occurred on line 98-80, with the real power flow estimated as 0.975MW using the MCS method, this implies the UT+UDR evaluated the power flow as 1.099MW while the PEM see it as 3.267MW; with the relative errors been 0.124MW and 2.292MW respectively. It is clear that the error using the PEM is significant and can impact negatively by oversizing the cable thus resulting into higher overall cost.

On the contrary, underestimation of the power flow is noticed on line 80-97 with the MCS giving 2.506MW while the UT+UDR and PEM gave 2.372MW and 0.151MW respectively. This amounts to a relative error of -0.134MW and -2.355MW respectively. This significant underestimation using the PEM can result into serious overflow in the line if the design was based on the method. It is observed that the errors due to the UT+UDR are generally lower; overall the maximum relative error using the UT+UDR amounted to 0.627MW while that for PEM was 8.42MW which is very significant.

8.7. Précis

In most real life problems, with power system not being an exemption, a dependence pattern is often noticed amongst parameters due to similarity in factors affecting them. For instance, in power systems, customers located within the same region may have similar load demand due to the prevailing weather condition. This spatial dependence is also evident in wind turbines located within the same area since similar wind speed reaches them. Therefore, for a method to be adaptable for load flow studies, it must be able to accommodate the effects of dependence amongst the uncertain variables.

In this chapter, the enhanced unscented transform method from the previous chapter is modified to incorporate the effect of linear dependence (correlation). The transformation method based on Cholesky decomposition have been proposed for incorporating correlation effects using the UT+UDR method while the Nataf transformation was employed for the Monte Carlo Simulation involving non-Gaussian variables. The performance of the Cholesky decomposition based UT+UDR was evaluated using three cases to ensure the adaptability of the technique for both Gaussian and non-Gaussian randomly distributed and correlated variables in either small or large systems. In all the cases, the performance of the UT+UDR was compared with those from the MCS and the PEM to rightly place the proposed method amongst existing techniques.

In cases where all the variables follow the Gaussian distribution and correlation is neglected, the UT+UDR and the PEM give approximately the same level of

accuracy. However, when correlation amongst the variables is considered, the UT+UDR outperform the PEM. For both cases, errors in the average values were generally lower while the effect of correlation is much more visible in the standard deviation. From the results generally, it is observed that the effects of correlation on the output parameters can be significant and as such should not be neglected. For the correlation between the wind farms where the output power follows non-Gaussian distribution, the PEM gives a wide range of errors due to its poor performance in representing arbitrary distributions. Overall, results obtained attest to the accuracy of the UT+UDR method with Cholesky decomposition (added to incorporate correlation) as compared with the PEM. Also, the proposed method functions well for both Gaussian and Non-Gaussian distributed random variables in both small and large systems as demonstrated in this chapter.

8.8. References

[8.1] G. Papaefthymiou and D. Kurowicka, "Using Copulas for Modeling Stochastic Dependence in Power System Uncertainty Analysis," in *IEEE Trans. on Power Systems,* vol. 24, no. 1, pp. 40-49, Feb. 2009

[8.2] P. Jackel, "A note on multivariate Gauss-Hermite quadrature," [Online]. Available: http://www.pjaeckel.webspace.virginmedia.com/ANoteOnMultivariateGaussHermiteQuadrature.pdf accessed [24/12/2012]

[8.3] R. E. Melchers, *Structural Reliability Analysis and Prediction*, 2nd ed., New York: John Wiley and Sons Ltd., 1999, pp. 374-385.

[8.4] S. Rahman, "Extended polynomial dimensional decomposition for arbitrary probability distributions," *ASCE Journal of Engineering Mechanics,* pp. 1439-1451, December 2009.

[8.5] M. S. Eldred, "Design under uncertainty employing stochastic expansion methods," *International Journal for Uncertainty Quantification,* vol. 1, no 2, pp. 119-146, 2011.

[8.6] M. S. Eldred and C. G. Webster, "Evaluation of Non-Intrusive Approaches for Wiener-Askey Generalized Polynomial Chaos"

American Institute of Aeronautics and Astronautics, paper 2008-1892, pp.1-22

[8.7] J. E. Hurtado, *Structural Reliability*, Berlin Heidelberg: Springer-Verlag, 2004, pp. 1-43.

[8.8] I. Lee, K.K. Choi, L. Du and D. Gorsich, "Dimension reduction method for reliability-based robust optimization," *Journal of Computers and Structure*s, vol. 86pp. 1550-1562, 2008.

[8.9] Y. Noh, K. K. Choi and L. Du, "Reliability-based design optimization of problems with correlated input variables using a Gaussian Copula," *Springer Journal of Structural Multidisciplinary Optimization*, vol. 38, pp. 1-16, 2008.

[8.10] B. Huang and X. Du, "Uncertainty Analysis by dimension reduction integration and saddlepoint approximations," Journal of Mechanical Design, pp 26-33, vol. 128, Jan 2006

[8.11] H. Li, Z. Lu and X. Yuan, "Nataf transformation based point estimate method," *Chinese Science Bulletin*, Springer, vol. 53, no. 17, pp. 2586-2592, Sept. 2008.

[8.12] M. Rosenblatt, "Remarks on a multivariate transformation," *The Annals of Mathematical Statistics*, vol. 23, no. 3, pp. 470-472, 1952.

[8.13] J.M. Morales, L. Baringo, A.J. Conejo and R. Minguez, "Probabilistic Power Flow with Correlated Wind Sources," *IET Generation Transmission and Distribution*, vol. 4, iss. 5, pp. 641-651, 2010

Chapter 9

Conclusion

This chapter summaries the findings from the work and examines how the objectives of the research have been met. Areas which can be further researched into are highlighted in the last section.

9.1. Load Flow Analysis in Modern Day Power Systems

Power systems have gradually evolved to reflect the need for sustainable solutions while meeting up the soaring demands of modern day consumers. The global call for a reduction in carbon emissions and the gradual depletion of fossils have put renewable energy (RE) generation as one major way of meeting up with this increasing demands with reduced emission potentials. This alongside the deregulation of the electricity market has helped increase the penetration level of RE systems within the modern day power grid as discussed in Chapter 2.

Considering the huge support for RE systems and the various structures (incentives and regulations) put in place in various countries of the World, more of these systems are expected to come on-stream. With this level of penetration, their effects cannot be trivialized bearing in mind that their outputs are mostly uncertain.

Load flow analysis is the most popular analysis carried out to study the performance of the power system. An analysis looking into the various techniques for load flow studies in both transmission (Newton Raphson, Gauss Seidel etc) and distribution (Backward/forward sweep, Newton method) was given in Chapter 2 alongside the mathematical formulation for the Newton Raphson and Backward/Forward sweep methods. These conventional load flow methods are deterministic in nature, hence the concept of probabilistic load flow studies was introduced as a means of accounting for uncertainties within the power systems.

The traditional way for carrying out probabilistic load flow is based on the Monte Carlo Simulation method. Though the method proves to be very accurate, its computation becomes burdensome and time consuming as the system size increases which necessitated the search for alternative methods able to give a good balance between speed and accuracy. A detailed evaluation of some of the initially proposed methods was presented in Chapter 4. Considering the reviewed methods, the features of an effective probabilistic load flow method were identified and reiterated below;

i. Minimal computational time and burden

ii. Good level of accuracy comparable with the MCS method

iii. Applicable to practical small and large systems

iv. Able to accurately model uncertainties within the network irrespective of their distribution

v. Able to represent the dependency that exists between random uncertainties in the network.

These informed the objective of this work which was to develop a fast and accurate technique for carrying out probabilistic load flow studies.

9.2. The Unscented Transform

The unscented transform (UT) has been successfully applied to non-linear electromagnetic compatibility statistical problems and as such seen as a viable method for probabilistic load flow studies. The method is based on discretising the continuous probability density function of a variable using deterministically chosen points referred to as sigma points having corresponding weights such that the moments of both the continuous and discrete points are the same.

Since the continuous distribution of uncertain variables in power systems can be characterised based on previous data set as explained in Chapter 3, the UT method was explored for the load flow problem in Chapter 5. The results obtained using the UT method for univariate problems with Gaussian

distribution show that its accuracy improves as the number of sigma points increases. The method can be extended to treat multivariate problems using either Taylor series multivariate expansion (T.S.M) or a general set which views the problem in the Euclidean space and thus uses the edges of an *n*-dimensional cube. The first method resulted in solving huge simultaneous equations which makes it impractical for systems with large number of uncertainties. Similarly, the sigma points for the general set increases as the number of uncertain variables increases. For problems involving more than 20 random (uncertain) variables, the techniques become even more computationally burdensome than the MCS thus defeating the objective.

Another aspect considered is the performance of the UT method for variables whose distribution does not follow the Gaussian distribution. This is very important since the wind speed and power distributions follow the Weibull and a composite distribution (truncated Weibull and discrete distributions) respectively as already validated in Chapter 3 and Appendix B. From the results, it was observed that a minimum of 5 sigma points are needed in discretizing each random variable for the wind power distribution to obtain a reasonable level of accuracy. This is because only 3 sigma points are chosen in the continuous part of the wind power distribution. For multivariate problems, the method resulted in negative unrealistic sigma points which make the two techniques (Taylor series multivariate expansion and the general set) inapplicable.

This implies the conventional Unscented Transform method is inapplicable to the probabilistic load flow problem since practical power systems often comprise of several uncertain distributions.

9.3. Enhanced Unscented Transform for Load Flow Studies

In view of the shortfalls of the conventional Unscented Transform methods, an alternative way of viewing UT as a Gaussian quadrature problem was discussed in Chapter 6. With this, the sigma points are chosen as the roots of a polynomial orthogonal to the probability density function of the variable whose sigma points are desired. Several distributions such as the Gaussian have

known orthogonal polynomial (Hermite Polynomial) which makes this method easy to adapt. Although, to be able to apply this technique to the wind speed and power distribution (and other arbitrary distributions), their orthogonal polynomials have to be built. The Stieltjes procedure was employed as a robust approach for building the coefficient of the orthogonal polynomial using a discretisation approach. The sigma points are then estimated from a tridiagonal Jacobian matrix while weights are easily estimated from the eigenvector. This method is seen to be very accurate and applicable to any variable whose distribution function is known.

Tensor product was used in extending the method to multivariate problems in the latter part of Chapter 6. The method simply implies carrying out *n*-dimensional integration for *n* independent variables. The method proved accurate when compared with the MCS method and also outperforms other methods considered (Point Estimate Method (PEM) and Cumulant Method). However, it suffers from the *curse of dimensionality* problem since the number of sigma points grow very rapidly as the number of uncertain parameters within the system increases. This implies that only two (that is; ii, iv) of the five features described in section 9.1 were resolved.

The dimension reduction technique was proposed in Chapter 7 to drastically reduce the number of sigma points required for multivariate problems. With the dimension reduction method, the multivariate problem is decomposed into a series of *n*-dimensional functions thus reducing the problem to a *D*-dimensional integration unlike the Tensor product. The univariate dimension reduction (UDR) and bivariate dimension reduction (BDR) with D=1 and 2 respectively were explored. With the UDR, the Taylor series expansion of *n*-dimensional function is terminated after the first order while for BRD, after the second. This implies the latter requires more estimation points than the former, though it provides a better level of accuracy especially for higher order moments.

The practicality of the methods was demonstrated using a modified IEEE 14-bus test system with 24 random variables (discrete, Gaussian and wind power distributions) and a real distribution power system located in Samsun, North

coast of Turkey with 117 uncertain variables including wind power. Results obtained showed that both techniques (UDR and BDR) give similar level of accuracy for evaluating the mean and standard deviation while for higher order moments, the BDR performed better though at the expense of more computational time. Hence, the decision on which of the techniques to be employed for a particular scenario depends on the most pressing quantity between accuracy of higher order moments and speed.

In Chapter 8, the Cholesky decomposition technique was incorporated with dimension reduction to tackle the problem of linear dependency (correlation) which is common in power systems. The effect of correlation can be included in the UDR method once the correlation matrix is known. This correlation matrix is decomposed using the Cholesky decomposition technique with the lower triangular matrix employed to multiply the already defined sigma points. The method simply transforms the original sigma points without changing their initial weights.

Correlation is seen to affect the higher order moment and not the mean of a variable. Results obtained using 6, 14 and 118-bus test systems showed that the scheme is applicable to problems involving both Gaussian and non-Gaussian distributions, and gives a good level of accuracy (better than the PEM) when benchmarked using the MCS for all the cases considered. The scheme was only extended to the UDR method because of its lower computational requirement, also, it can easily be extended to the BDR method.

With this, the five criteria discussed have been satisfactorily met.

9.4. Summary of Contributions

The summary of the contributions of this work are given below;

1. The application of the Unscented Transform method to probabilistic load flow studies.
2. Representation of the wind power distribution as a composite distribution consisting of the truncated Weibull and discrete

distributions (at zero and rated powers). The validation of the cubic model for modelling the wind power.

3. Extension of the Unscented Transform method to solve problems involving arbitrary distributions using the Stieltjes procedure.
4. Enhancement of the Unscented Transform method to reduce the problem of *curse of dimensionality* based on dimension reduction techniques.
5. Incorporation of correlations (linear dependence) into univariate dimension reduction technique.

9.5. Further Research Areas

The suggestions for further works have been sectionalized into two. The first centres on areas related to probabilistic load flow methods in general while the last part focuses on possible areas for future works in terms of the Unscented Transform method.

9.5.1. Probabilistic Load Flow and Modern Power Systems

The wind power system was the main focus of the work since it is about the largest single unit renewable energy system (aside from hydropower) that can be installed. It will be interesting to check the combined effect of several other renewable energy generators on the power system so as to understand the interaction among them. This alongside analysing real data will help create the true picture of the emerging modern day power system.

In this work, like in most of the previous researches, dependency was viewed as linear dependence (correlation), which is not always true as explained in Chapter 3. Since zero correlation does not necessarily imply zero dependence, it will be necessary to explore other forms of dependence that may exists between component variables of the power system as this will assist in proper prediction and planning.

The output of a wind power system is dependent not only on the wind speed but also on the direction of the wind. Having a single wind speed and or power model able to take the effect of directional wind into account will produce better estimate of the output wind power from a wind farm.

9.5.2. Unscented Transform

From this work, several improvements have been proposed to the conventional Unscented Transform which are; the Stieljes procedure for determining the sigma points of arbitrary distributions, dimension reduction to overcome the *curse of dimensionality* and transformation (using Cholesky decomposition) to incorporate correlation. The feasibility of employing other techniques needs to be explored. For instance, dimension reduction can be carried out using Karhunen-Loève transforms (principal component analysis) or the independent component analysis amongst other methods [9.1].

Another area that can be explored is incorporating the effect of correlation using multivariate orthogonal polynomial. This will imply accounting for the effect of correlation at the initial stage of solving for the sigma points.

9.6 References

[9.1] K. Fodor, "A survey of dimension reduction techniques," [Online]. Available: https://computation.llnl.gov/casc/sapphire/pubs/148494.pdf [Accessed 03/11/2012].

Appendix A

Transformer Parameters [A.1]

The parameters for a grounded Wye to grounded Wye (Y-Y) 3 phase connected transformer as discussed in Section 2.3.2.1 of Chapter 2 are given below

$$c_t = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad d_t = \frac{1}{n_t}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$A = \frac{1}{n_t}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} Z_a & 0 & 0 \\ 0 & Z_b & 0 \\ 0 & 0 & Z_c \end{bmatrix}$$

The parameters for a Delta to grounded Wye (Δ-Y) 3 phase connected transformer are given below.

$$c_t = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad d_t = \frac{1}{n_t}\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}$$

$$A = \frac{1}{n_t}\begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} Z_a & 0 & 0 \\ 0 & Z_b & 0 \\ 0 & 0 & Z_c \end{bmatrix}$$

where n_t is the transformer turns ratio and Z_a, Z_b and Z_c are the per-unit impedance of the transformer windings on the load side of the transformer.

Reference

[A.1] W. H. Kersting, Distribution System Modelling, CRC Press, LLC 2006

Appendix B

Wind Speed and Power Analysis

B.1. Location Description

The wind speed is from a proposed wind farm site within the Ondokuz Mayis University campus, Samsun North coast of Turkey ($41^{o}22'N$ $36^{o}12'E$). The measurement was carried out over a 12month period (August 2009-July 2010) with wind speed measurements every 10 minutes at 40m, 50m and 60m above the ground level. As expected the best wind speed were observed at 60m height and will be used in this analysis.

A summary of the average and maximum wind speed for each month from the raw wind speed data is shown in Table B.1. The average measured air density for the area is $1.557kg/m^3$.

Table B. 1: Average and Maximum Wind Speed for Samsun

Month	Average Wind Speed (m/s)	Maximum Wind Speed (m/s)
August	5.0	11.92
September	5.3	19.48
October	4.8	24.22
November	5.1	34.48
December	6.3	36.40
January	6.2	31.51
February	7.1	31.40
March	5.6	25.41
April	4.9	19.73
May	4.3	20.67
June	5.4	21.17
July	5.4	16.91

B.2. Wind Data Fitting: Criteria

The wind data was analysed using the EasyFit analysis tool [B.1]. To ensure the data is rightly fitted, the three goodness of fit test available on EasyFit are used as the determining factor for choosing the best distribution. The three tests namely the Kolmogorov-Smirnov, Anderson-Darling and the Chi-Square tests are briefly described below.

B.2.1. The Kolmogorov-Smirnov (K-S) Test

The K-S test is a non-parametric test used to determine if a set of samples come from a hypothesized continuous distribution. The K-S test is based on the empirical cumulative distribution function (ECDF) and it measures the similarity by quantifying the distance between the empirical distribution of the sample set with those for the cumulative distribution function of the reference distribution [B.1-B.3].

One key advantage of the K-S test is that it does not depend on the cumulative distribution function been tested, also, it is an exact test [B.2]. However, it is not applicable to non-continuous distributions, it is also more sensitive near the centre than at the tail ends thus limiting its accuracy in distributions with significant tail features [B.2].

The largest difference between the empirical and theoretical distributions D is obtained using (B.1) [B.1, B.2].

$$D = \max_{1 \le i \le N} \left(F(Y_i) - \frac{i-1}{N}, \frac{i}{N} - F(Y_i) \right) \qquad (B.1)$$

B.2.2. Anderson-Darling Test

The Anderson-Darling test is a modification to the K-S test. Like the K-S test, it is used to determine if a set of samples come from a specified distribution but, it gives more weight to the tail as compared with the K-S test which makes it a more sensitive test [B.1, B.4].

$$A^2 = -N - \sum_{i=1}^{N} \frac{(2i-1)}{N}\left[\ln F(Y_i) + \ln(1 - F(Y_{N-i+1}))\right] \qquad (B.2)$$

B.2.3. The Chi-Square Test (χ^2)

The Chi-Square test is used to determine if a sample comes from a population with specific distribution [B.1, B.5]. It is applied to binned distributions and also to continuous distributions like the K-S test. The value of the statistics is dependent on how the data is binned. For optimal bin size, EasyFit employs (B.3) to get *k* which is the optimal number of bins for a sample size *N*. The validity and accuracy of the test also depends on the sample size [B.5].

$$k = 1 + \log_2 N \qquad (B.3)$$

The Chi-Square statistics is defined by (B.4).

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \qquad (B.4)$$

$$E_i = F(x_2) - F(x_1) \qquad (B.5)$$

where O_i is the observed frequency for bin *i*, E_i is the expected frequency for bin *i* defined by (B.5) and *F* is the cumulative distribution function of the probability distribution under test, while x_1 and x_2 are the limits for bin *i*.

A detailed treatment of the various goodness of fit test can be found in [B.1-B.5]

B.3. Wind Data Fitting: Result

Several distributions such the Weibull distribution [B.6, B.7], normal distribution [B.8], Gamma distribution [B.9], the Rayleigh distribution [B.10] and the lognormal distribution have been used in modelling wind speed. In this analysis, six distributions; normal, exponential, lognormal, Rayleigh, Gamma and Weibull distributions are tested to determine which of them gives the best fit to the data.

The histogram for the wind data for the whole year alongside the probability density function curves using each of the 6 distribution functions is given in Fig. B.1. The statistics using the three criteria earlier discussed in section B.2 above are presented in Table B.2.

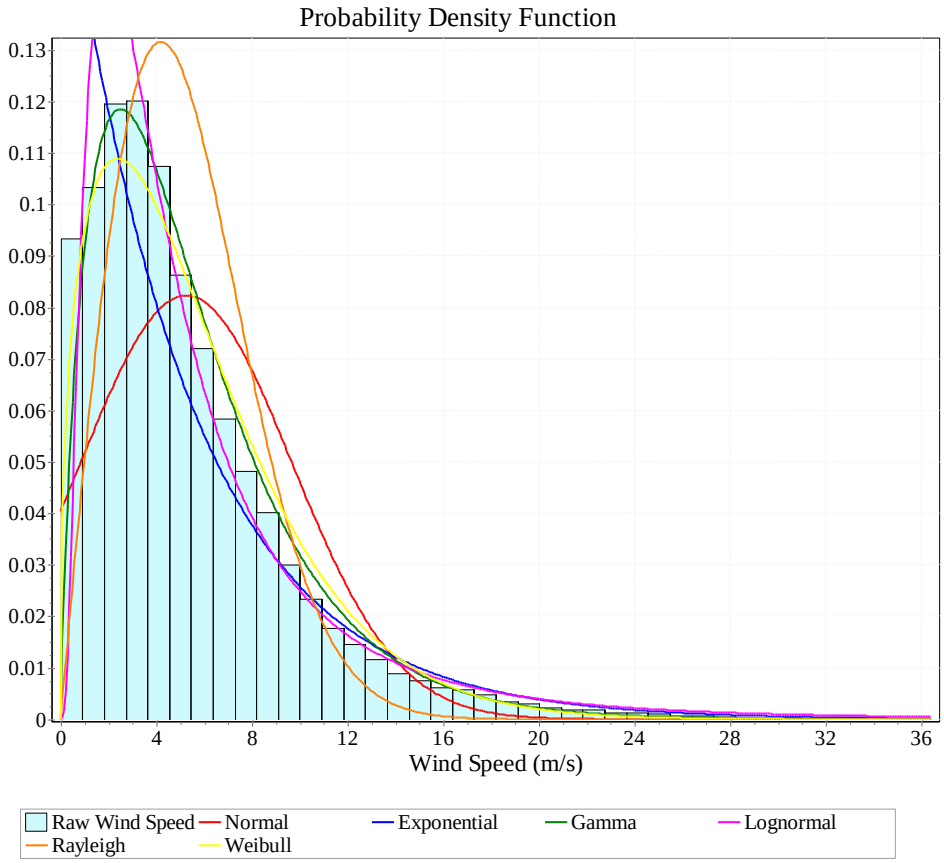

Fig B. 1: Histogram for the Raw Wind Speed Data and PDF Curves Using 6 Distributions

Table B.2: Statistics for the Goodness of Fit

Distribution	Goodness of Fit Test		
	Kolmogorov-Smirnov	Anderson-Darling	Chi Square
Normal	0.11665	1415.3	8926.1
Exponential	0.09742	5048.7	4490.7
Gamma	0.06022	4709.0	1423.0
Lognormal	0.06403	5147.6	2848.1
Rayleigh	0.12636	6859.4	12316
Weibull	0.05993	4618.9	1182.0

From the results, the commonly used Weibull distribution is seen to give the best result in terms of the K-S test criterion. To further affirm the suitability of the Weibull distribution for the wind speed data for the location considered, the statistics for the other test are assessed. The Weibull distribution is seen to be the second best considering the Anderson-Darling test while it is the best of them based on the Chi Square test. The Weibull distribution is therefore assumed to be the best fit for the wind data and used in line with some other research work [B.6, B.7]. It is worth mentioning that the same critical value is used for all the distributions in the EasyFit software [B.1].

The shape and scale parameter for the Weibull distribution for each month considered and the final one for the whole 12 month period are presented in Table B.3.

Table B.3: Weibull Parameter for the Wind Speed

Month	Shape Parameter	Scale Parameter
August	1.8208	5.5278
September	1.6239	5.7843
October	1.5575	5.6523
November	1.1462	5.8145
December	1.2298	6.7587
January	1.1482	7.6337
February	1.2958	7.5652
March	1.4534	6.1959
April	1.7200	5.5066
May	1.5433	4.7191
June	1.3664	5.9271
July	1.4513	5.6241
Whole Year	1.3770	6.1345

B.4. Wind Power Model

Using the wind speed distribution data obtained above, the output wind power can be easily determined. The appropriateness of the two models (cubic and linear models) previous used in the literature in representing the wind power model as discussed in Section 3.5.4 (Chapter 3), is appraised. Since the measured output wind power is not available, the output power obtained by using the raw wind speed data and the lookup table from the turbine manufacturer is assumed to be the benchmark. The EWT500 (500kW) wind turbine [B.11] is assumed to be installed on the farm since wind installations over 500kW require planning permission from the Turkish Government [B.12].

The mean output wind power for the 12 month period is analysed using 3 principal techniques described below.

1. The raw wind speed data and the lookup table generated from the wind turbine characteristics curve. (Benchmark)
2. The wind speed data generated using the Weibull parameters for both the Monte Carlo Simulation and Unscented Transform methods and then combined with the linear model (B.6, B.7) or the cubic model (B.8, B.9) to obtain the final output wind power.

$$P_w = \begin{cases} 0 & v \le v_{ci} \quad \text{(Region A)} \\ Cv - D & v_{ci} < v \le v_r \quad \text{(Region B)} \\ P_r & v_r \le v \le v_{co} \quad \text{(Region C)} \\ 0 & v > v_{co} \quad \text{(Region D)} \end{cases} \tag{B.6}$$

$$C = \frac{P_r}{(v_r - v_{ci})}, \qquad D = Cv_{ci} \ \text{(linear)} \tag{B.7}$$

$$P_w = \begin{cases} 0 & v \le v_{ci} \quad \text{(Region A)} \\ Cv^3 - D & v_{ci} < v \le v_r \quad \text{(Region B)} \\ P_r & v_r \le v \le v_{co} \quad \text{(Region C)} \\ 0 & v > v_{co} \quad \text{(Region D)} \end{cases} \tag{B.8}$$

$$C = \frac{P_r}{(v_r^3 - v_{ci}^3)}, \qquad D = Cv_{ci}^3 \ \text{(cubic)} \tag{B.9}$$

3. The output wind power generated directly using the wind power models based on both the linear (B.10) and cubic models (B.11, B.12) for the Unscented Transform method.

$$f(P_w)=\begin{cases}1+\exp\left[-\left(\frac{v_{co}}{\beta}\right)^{\alpha}\right]-\exp\left[-\left(\frac{v_{ci}}{\beta}\right)^{\alpha}\right] & P_w=0\\ \frac{\alpha}{C\beta}\left(\frac{P_w+D}{\beta C}\right)^{\alpha-1}\exp\left[-\left(\frac{P_w+D}{\beta C}\right)^{\alpha}\right] & 0<P_w<P_r\\ \exp\left[-\left(\frac{v_r}{\beta}\right)^{\alpha}\right]-\exp\left[-\left(\frac{v_{co}}{\beta}\right)^{\alpha}\right] & P_w=P_r\end{cases} \quad \text{(B.10)}$$

$$f(P_w)=\begin{cases}1+\exp\left[-\left(\frac{v_{co}}{\beta}\right)^{\alpha}\right]-\exp\left[-\left(\frac{v_{ci}}{\beta}\right)^{\alpha}\right] & P_w=0\\ \frac{\alpha}{3C\beta^{\alpha}}\left(\frac{P_w+D}{C}\right)^{(\alpha/3)-1}\exp\left[-\frac{((P_w+D)/C)^{\alpha/3}}{\beta^{\alpha}}\right] & 0<P_w<P_r\\ \exp\left[-\left(\frac{v_r}{\beta}\right)^{\alpha}\right]-\exp\left[-\left(\frac{v_{co}}{\beta}\right)^{\alpha}\right] & P_w=P_r\end{cases} \quad \text{(B.11)}$$

$$A=\frac{P_w+D}{C} \quad \text{(B.12)}$$

The results using the various schemes are presented in Table B.4.

Table B.4: Mean Output Power for a 500kW Wind Turbine

Raw Data (kW)	MCS Linear (kW)	MCS+ Cubic (kW)	UT Linear (kW)	UT Cubic (kW)	UT Pwr Linear (kW)	UT Pwr Cubic (kW)
141.272	194.190	141.148	195.313	136.350	193.936	141.056

Legend

Raw Data: Results obtained using the measured data and the lookup table.

MCS+ Linear: Results obtained by generating the wind speed using the Weibull parameter obtained in a random number generator and then combining this wind speed with the linear model in equation (B.6, B.7).

MCS+ Cubic: Results obtained by generating the wind speed using the Weibull parameter obtained in a random number generator and then combining this wind speed with the cubic model in equation (B.8, B.9).

UT+ Linear: Results obtained by generating the wind speed using the Weibull parameter obtained based on the UT method and then combining the sigma points and weights of the wind speed with the linear model in equation (B.6, B.7).

UT+ Cubic: Results obtained by generating the wind speed using the Weibull parameter obtained based on the UT method and then combining the sigma points and weights of the wind speed with the cubic model in equation (B.8, B.9).

UT+ Linear (Pwr): Results obtained by generating the output wind power directly using (B.10).

UT+ Cubic (Pwr): Results obtained by generating the output wind power directly using (B.11, B.12).

From the results, it is seen that all methods based on the cubic model generally give better estimate of the output power than linear model. In fact, the linear model overestimates the output power from the wind turbine. For the UT method, it is clear that using the wind power model directly gives a better estimate than using the sigma points and weights generated using the wind speed. The result also further confirms the accuracy of the UT method as it gives a close match with the real data and the Monte Carlo Simulation method.

B.5. Further Analysis Using the IEEE 14-Bus Test System

To further appraise the performance of the various techniques and to validate the performance of the cubic model using the UT method, a simple load flow analysis is carried out using the IEEE 14-bus test system. A group of 100, 500kW wind turbine is assumed to be located on bus 9 of the test system with maximum output of 50MW. Load flow analysis was then carried out using the techniques discussed in section B.4. above for generating the output wind power. The results for the output mean voltage a magnitude and active power on all the buses is presented in Table B.5 and B.6 respectively. To further reflect the performance of the methods, the average errors in the first four moments for the voltage magnitude and active power flows are shown in Tables B.7 and B.8.

Table B.5: Mean Voltage Magnitude Using the Various Techniques

Bus	Raw Data	MCS Cubic	MCS Linear	UT+ Cubic	UT+ Linear	UT+ Cubic (Pwr)	UT Linear (Pwr)
1	1.06	1.06	1.06	1.06	1.06	1.06	1.06

2	1.045	1.045	1.045	1.045	1.045	1.045	1.045
3	1.01	1.01	1.01	1.01	1.01	1.01	1.01
4	1.0206	1.0208	1.0196	1.0207	1.0196	1.0206	1.0196
5	1.0220	1.02212	1.0212	1.0221	1.0211	1.0220	1.0211
6	1.07	1.07	1.07	1.07	1.07	1.07	1.07
7	1.0641	1.0643	1.0633	1.0642	1.0632	1.0642	1.0632
8	1.09	1.09	1.09	1.09	1.09	1.09	1.09
9	1.0596	1.0598	1.0584	1.0597	1.0583	1.0596	1.0583
10	1.0541	1.0543	1.0531	1.0542	1.0530	1.0542	1.0530
11	1.0587	1.0588	1.0581	1.0587	1.0580	1.0587	1.0580
12	1.0553	1.0554	1.0553	1.0553	1.0553	1.0553	1.0553
13	1.0510	1.0511	1.0508	1.0511	1.0508	1.0510	1.0508
14	1.0379	1.0381	1.0371	1.0380	1.0371	1.0379	1.0371

Table B.6: Active Power Flow Using the Various Techniques

Line No	Raw Data	MCS Cubic	MCS Linear	UT+ Cubic	UT+ Linear	UT+ Cubic (Pwr)	UT Linear (Pwr)
1-2	-142.05	-142.02	-145.57	-141.71	-145.75	-142.03	-145.66
1-5	-67.343	-67.328	-69.147	-67.171	-69.241	-67.336	-69.193
2-3	-68.713	-68.707	-69.453	-68.644	-69.490	-68.710	-69.472
2-4	-50.096	-50.084	-51.554	-49.957	-51.630	-50.090	-51.591
2-5	-37.199	-37.190	-38.343	-37.091	-38.402	-37.194	-38.373
3-4	25.931	25.938	25.169	26.003	25.130	25.935	25.149
4-5	57.337	57.325	58.779	57.197	58.856	57.331	58.815
4-7	-21.139	-21.120	-23.455	-20.917	-23.577	-21.129	-23.514
4-9	-12.114	-12.103	-13.437	-11.987	-13.507	-12.108	-13.471
5-6	-39.605	-39.593	-41.112	-39.464	-41.187	-39.599	-41.151
6-11	-4.5847	-4.5772	-5.4944	-4.4987	-5.5407	-4.5810	-5.5178
6-12	-7.3725	-7.372	-7.4877	-7.3618	-7.4933	-7.3720	-7.4907
6-13	-16.155	-16.151	-16.618	-16.111	-16.642	-16.153	-16.631
7-8	3.4E-15	2.8E-15	2.9E-15	1.5E-14	1.4E-15	2.5E-15	4.9E-15
7-9	-21.139	-21.120	-23.455	-20.917	-23.577	-21.129	-23.514

9-10	-7.9269	-7.9344	-7.0192	-8.0132	-6.9723	-7.9307	-6.9959
9-14	-11.015	-11.020	-10.444	-11.069	-10.415	-11.018	-10.429
10-11	1.0847	1.0772	1.9944	0.9987	2.0407	1.0809	2.0178
12-13	-1.2674	-1.2665	-1.3821	-1.2567	-1.3877	-1.2670	-1.3851
13-14	-3.8848	-3.8801	-4.4564	-3.8308	-4.4855	-3.8825	-4.4711

Table B.7: Average Percentage Error in the First Four Moments of the Voltage Magnitude

Moment	MCS Cubic [%]	MCS Linear [%]	UT+ Cubic [%]	UT+ Linear [%]	UT+ Cubic (Pwr) [%]	UT Linear (Pwr) [%]
Mean	0.0077	0.0457	0.0039	0.0488	1.46e-4	0.0485
Std	1.8966	0.8538	2.7732	1.5359	0.2431	1.2379
Skew	8.1939	29.792	13.642	20.327	1.2383	31.160
Kurt	10.161	18.969	8.4696	6.2769	1.4873	20.250

Table B.8: Average Percentage Error in the First Four Moments of the Voltage Magnitude

Moment	MCS Cubic	MCS Linear	UT+ Cubic	UT+ Linear	UT+ Cubic (Pwr)	UT Linear (Pwr)
Mean	0.0835	10.143	0.9608	10.662	0.0419	10.404
Std	0.0076	0.8895	1.9283	2.2366	0.1801	1.3062
Skew	2.0682	36.639	18.459	28.623	1.2690	37.551
Kurt	3.3671	28.638	17.394	9.1160	2.1495	29.632

From the tables, it is clear that the UT+ Cubic (Pwr) gives a better level of accuracy with reference to the raw data (benchmark) in almost all the cases than the other techniques considered including the MCS method. This informed its choice for simulations in this work.

B.6. References

[B.1] EasyFit© User Manual 2008

[B.2] NIST/SEMATECH e-Handbook of Statistical Methods, Kolmogorov-Smirnov Goodness of Fit Test. Available online : http://www.itl.nist.gov/div898/handbook/eda/section3/eda35g.htm [Accessed 25/01/2013]

[B.3] Kolmogorov-Smirnov Test. Encyclopedia of Mathematics. Available online: http://www.encyclopediaofmath.org/index.php?title=Kolmogorov%E2%80%93Smirnov_test&oldid=22659 [Accessed 25/01/2013]

[B.4] NIST/SEMATECH e-Handbook of Statistical Methods, Anderson-Darling Test. Available online : http://www.itl.nist.gov/div898/handbook/eda/section3/eda35e.htm [Accessed 25/01/2013]

[B.5] NIST/SEMATECH e-Handbook of Statistical Methods, Chi-Square Goodness of fit. Available online : http://www.itl.nist.gov/div898/handbook/eda/section3/eda35f.htm [Accessed 25/01/2013]

[B.6] D.M. Outcalt. "Probabilistic Load Flow for High Wind Penetrated Power Systems based on a Five Point Estimation Method," PhD Thesis, University of Wisconsin-Milwaukee, December 2009.

[B.7] J. M. Morales, L. Baringo, A. J. Conejo and R. Minguez, "Probabilistic Power Flow with correlated wind sources" IET Gener. Transm. Distrib., vol.4, Iss. 5, pp641-651, 2010

[B.8] M. Aien, M. Fotuhi-Firuzabad and F. Aminifar, "Probabilistic Load Flow in Correlated Uncertain Environment Using Unscented Transformation", *IEEE Trans. Power Systems*, vol.27, no. 4., pp 2233-2241, Nov 2012

[B.9] J. Usaola "Probabilistic load flow with wind production uncertainty using cumulants and Cornish–Fisher expansion" [http://e-archivo.uc3m.es]

[B.10] S. Mathew, K. P. Pandey, A. V. Kumar, "Analysis of wind regimes for energy estimation", Renewable Energy, Vol. 25, Iss 3, pp. 381-399, March 2002

[B.11] EWT500, available online: http://www.genatec.co.uk/downloads/EWT.pdf [accessed 11/01/2012]

[B.12] H. Gedik, and U. S. Boz, "Renewable energy in Turkey: recent regulatory developments" EBRD, Law in Transition, Oct. 2011 Available online: www.ebrd.com/lawintransition

Appendix C

The Test Systems

The system data for the 6, 14 and 118 Test systems used in the work are presented in this appendix [C.1], [C.2]. The data for the 44 bus Samsun distribution test system is also presented.

C.1. IEEE 6 Bus Test System [C.1], [C.2]

The single line diagram and the bus, branch and generator parameters in presented below.

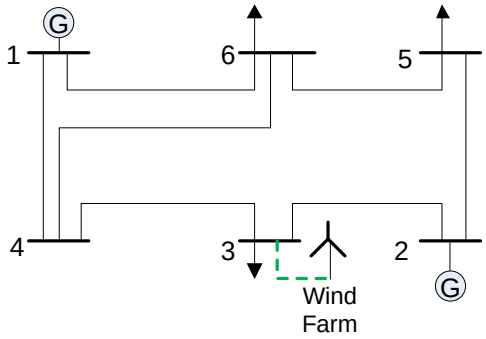

Fig C.1: Single Line Diagram for the 6 Bus Test System Modified to Include the Wind Farm

C.1.1. Generator Data for 6 Bus Test System

Table C.1: Generator Parameter for the 6 Bus Test System

Bus No	Active Power (MW)	Reactive Power (MVar)
1	0	0
2	45	24

C.1.1. Load Data for 6 Bus Test System

Table C.2: Load Data for the 6 Bus Test System

Bus No	Active Load (MW)	Reactive Load (MVar)	Voltage Mag. (pu)	Voltage Angle
1	0	0	1.05	0
2	0	0	1.07	0
3	55	13	1	0
4	0	0	1	0
5	30	18	1	0
6	60	15	1	0

Note: The base voltage is 230kV.

C.1.3. Branch Data for 6 Bus Test System

The Branch data for the 6 bus system are shown below. Where R represents the branch resistance, X is the branch reactance and B the line charging. All values are in per unit based on a base MVA of 100.

Table C.3: Branch Data for the 6 Bus Test System

From	To	R	X	B
1	4	0.08	0.37	0.02
1	6	0.123	0.518	0.06
2	3	0.0723	1.05	0.044
2	5	0.282	0.064	0.24
3	4	0	0.133	0.66
4	6	0.097	0.407	0.02
5	6	0	0.3	0.05

C.2. The IEEE 14 Bus Test System [C.1], [C.2]

The single line diagram and the bus, branch and generator parameters for the IEEE 14 bus test System are presented below. The location of the wind farm is also shown.

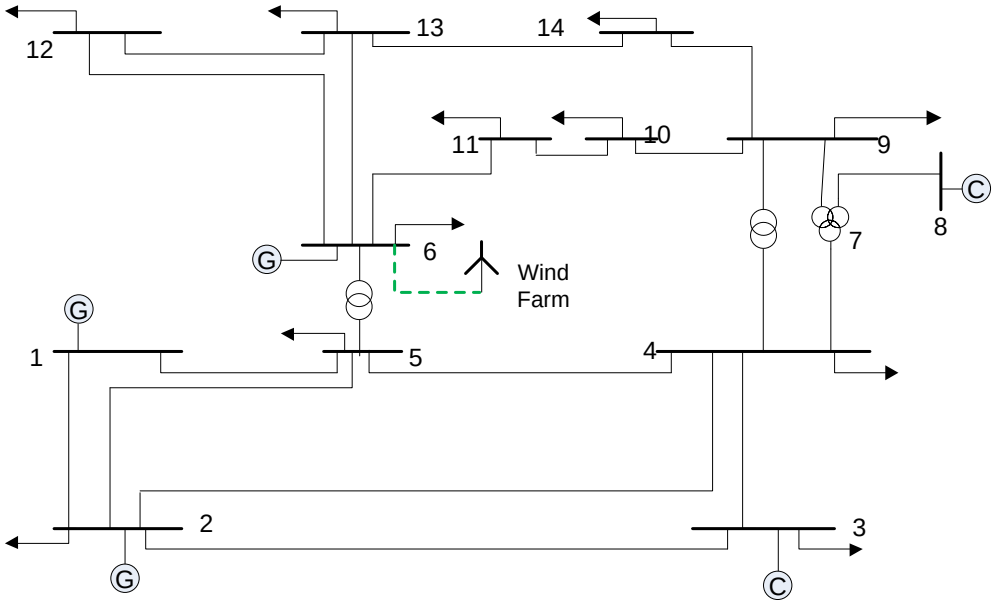

Fig C.2: The IEEE 14 Bus Test System Modified to Include Wind Farm

C.2.1. Generator Data For 14 Bus Test System

Table C.4: Generator Parameter for the IEEE 14 Bus Test System

Bus No	Active Power (MW)	Reactive Power (MVar)
1	232.4	-16.9
2	40.04	1.2
3	0	23.4
6	0	12.2
8	0	17.4

C.2.2. Branch Data for 14 Bus Test System

The Branch data for the IEEE 14 bus test system are shown below. Where R represents the branch resistance, X is the branch reactance and B the line charging. All values are in per unit based on a base MVA of 100.

Table C.5: Branch Data for the IEEE 14 Test System

From	To	R	X	B
1	2	0.01938	0.05917	0.0528
1	5	0.05403	0.22304	0.0492
2	3	0.04699	0.19797	0.0438
2	4	0.05811	0.17632	0.034
2	5	0.05695	0.17388	0.0346
3	4	0.06701	0.17103	0.0128
4	5	0.01335	0.04211	0
4	7	0	0.20912	0
4	9	0	0.55618	0
5	6	0	0.25202	0
6	11	0.09498	0.1989	0
6	12	0.12291	0.25581	0
6	13	0.06615	0.13027	0
7	8	0	0.17615	0
7	9	0	0.11001	0
9	10	0.03181	0.0845	0
9	14	0.12711	0.27038	0
10	11	0.08205	0.19207	0
12	13	0.22092	0.19988	0
13	14	0.17093	0.34802	0

C.2.3. Bus Data for 14 Bus Test System

Table C.6: Bus Data for the 14 Bus Test System

Bus No	Active Load (MW)	Reactive Load (MVar)	Voltage Mag. (pu)	Voltage Angle
1	0	0	1.06	0
2	21.7	12.7	1.045	-4.98
3	94.2	19	1.01	-12.72
4	47.8	-3.9	1	-10.33
5	7.6	1.6	1	-8.78
6	11.2	7.5	1.07	-14.22
7	0	0	1	-13.37
8	0	0	1.09	-13.36
9	29.5	16.6	1	-14.94
10	9	5.8	1	-15.1
11	3.5	1.8	1	-14.79
12	6.1	1.6	1	-15.07
13	13.5	5.8	1	-15.16
14	14.9	5	1	-16.04

C.3. 118 Test System

The single line diagram and the bus, branch and generator parameters for the IEEE 118 bus test System are presented below. The locations of the wind farms are also shown. The Base MVA is taken as 100 while the base voltage is 138kV.

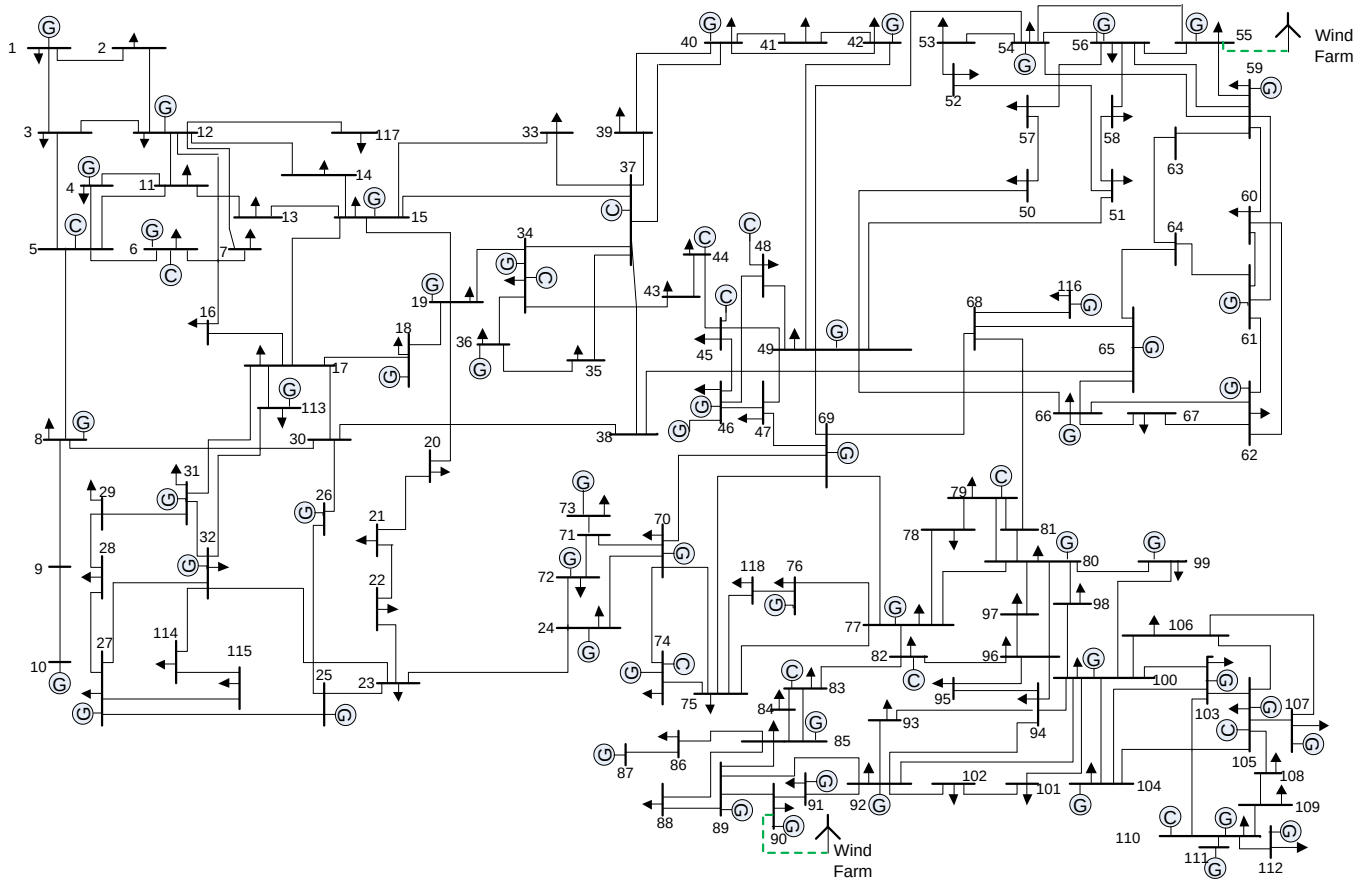

Fig C.3: Single Line Diagram of the IEEE 118 Test System

C.3.1. Bus Data for the IEEE 118 Bus Test System [C.1], [C.2]

Table C.7: Bus Data for The IEEE 118 Bus Test System

Bus No	Active Load (MW)	Reactive Load (MVar)	Voltage Mag. (pu)	Voltage Angle
1	51	27	0.955	10.67
2	20	9	0.971	11.22
3	39	10	0.968	11.56
4	39	12	0.998	15.28
5	0	0	1.002	15.73
6	52	22	0.99	13
7	19	2	0.989	12.56
8	28	0	1.015	20.77
9	0	0	1.043	28.02
10	0	0	1.05	35.61
11	70	23	0.985	12.72
12	47	10	0.99	12.2
13	34	16	0.968	11.35
14	14	1	0.984	11.5
15	90	30	0.97	11.23
16	25	10	0.984	11.91
17	11	3	0.995	13.74
18	60	34	0.973	11.53
19	45	25	0.963	11.05
20	18	3	0.958	11.93
21	14	8	0.959	13.52
22	10	5	0.97	16.08
23	7	3	1	21
24	13	0	0.992	20.89

25	0	0	1.05	27.93
26	0	0	1.015	29.71
27	71	13	0.968	15.35
28	17	7	0.962	13.62
29	24	4	0.963	12.63
30	0	0	0.968	18.79
31	43	27	0.967	12.75
32	59	23	0.964	14.8
33	23	9	0.972	10.63
34	59	26	0.986	11.3
35	33	9	0.981	10.87
36	31	17	0.98	10.87
37	0	0	0.992	11.77
38	0	0	0.962	16.91
39	27	11	0.97	8.41
40	66	23	0.97	7.35
41	37	10	0.967	6.92
42	96	23	0.985	8.53
43	18	7	0.978	11.28
44	16	8	0.985	13.82
45	53	22	0.987	15.67
46	28	10	1.005	18.49
47	34	0	1.017	20.73
48	20	11	1.021	19.93
49	87	30	1.025	20.94
50	17	4	1.001	18.9
51	17	8	0.967	16.28
52	18	5	0.957	15.32
53	23	11	0.946	14.35
54	113	32	0.955	15.26

55	63	22	0.952	14.97
56	84	18	0.954	15.16
57	12	3	0.971	16.36
58	12	3	0.959	15.51
59	277	113	0.985	19.37
60	78	3	0.993	23.15
61	0	0	0.995	24.04
62	77	14	0.998	23.43
63	0	0	0.969	22.75
64	0	0	0.984	24.52
65	0	0	1.005	27.65
66	39	18	1.05	27.48
67	28	7	1.02	24.84
68	0	0	1.003	27.55
69	0	0	1.035	30
70	66	20	0.984	22.58
71	0	0	0.987	22.15
72	12	0	0.98	20.98
73	6	0	0.991	21.94
74	68	27	0.958	21.64
75	47	11	0.967	22.91
76	68	36	0.943	21.77
77	61	28	1.006	26.72
78	71	26	1.003	26.42
79	39	32	1.009	26.72
80	130	26	1.04	28.96
81	0	0	0.997	28.1
82	54	27	0.989	27.24
83	20	10	0.985	28.42
84	11	7	0.98	30.95

85	24	15	0.985	32.51
86	21	10	0.987	31.14
87	0	0	1.015	31.4
88	48	10	0.987	35.64
89	0	0	1.005	39.69
90	163	42	0.985	33.29
91	10	0	0.98	33.31
92	65	10	0.993	33.8
93	12	7	0.987	30.79
94	30	16	0.991	28.64
95	42	31	0.981	27.67
96	38	15	0.993	27.51
97	15	9	1.011	27.88
98	34	8	1.024	27.4
99	42	0	1.01	27.04
100	37	18	1.017	28.03
101	22	15	0.993	29.61
102	5	3	0.991	32.3
103	23	16	1.001	24.44
104	38	25	0.971	21.69
105	31	26	0.965	20.57
106	43	16	0.962	20.32
107	50	12	0.952	17.53
108	2	1	0.967	19.38
109	8	3	0.967	18.93
110	39	30	0.973	18.09
111	0	0	0.98	19.74
112	68	13	0.975	14.99
113	6	0	0.993	13.74
114	8	3	0.96	14.46

115	22	7	0.96	14.46
116	184	0	1.005	27.12
117	20	8	0.974	10.67
118	33	15	0.949	21.92

C.3.2. Generator Data for the 118 Bus Test System

The data for the generators within the test system are presented in Table C.8 below. The maximum reactive powers for the generators are also presented alongside.

Table C.8: Generator Data for the IEEE 118 Bus Test System

Bus No	Active Power (MW)	Reactive Power (MVar)	Max. Reactive Power (MVar)
1	0	0	15
4	0	0	300
6	0	0	50
8	0	0	300
10	450	0	200
12	85	0	120
15	0	0	30
18	0	0	50
19	0	0	24
24	0	0	300
25	220	0	140
26	314	0	1000
27	0	0	300
31	7	0	300
32	0	0	42
34	0	0	24
36	0	0	24

40	0	0	300
42	0	0	300
46	19	0	100
49	204	0	210
54	48	0	300
55	0	0	23
56	0	0	15
59	155	0	180
61	160	0	300
62	0	0	20
65	391	0	200
66	392	0	200
69	516.4	0	300
70	0	0	32
72	0	0	100
73	0	0	100
74	0	0	9
76	0	0	23
77	0	0	70
80	477	0	280
85	0	0	23
87	4	0	1000
89	607	0	300
90	0	0	300
91	0	0	100
92	0	0	9
99	0	0	100
100	252	0	155
103	40	0	40
104	0	0	23
105	0	0	23

107	0	0	200
110	0	0	23
111	36	0	1000
112	0	0	1000
113	0	0	200
116	0	0	1000

C.3.3. Branch Data for the IEEE 118 Bus Test System

The Branch data for the IEEE 118 bus test system are shown below. Where R represents the branch resistance, X is the branch reactance and B the line charging. All values are in per unit based on a base MVA of 100.

Table C.9: Branch Data for the 118 Bus Test System

From	To	R	X	B
1	2	0.0303	0.0999	0.0254
1	3	0.0129	0.0424	0.01082
4	5	0.00176	0.00798	0.0021
3	5	0.0241	0.108	0.0284
5	6	0.0119	0.054	0.01426
6	7	0.00459	0.0208	0.0055
8	9	0.00244	0.0305	1.162
8	5	0	0.0267	0
9	10	0.00258	0.0322	1.23
4	11	0.0209	0.0688	0.01748
5	11	0.0203	0.0682	0.01738
11	12	0.00595	0.0196	0.00502
2	12	0.0187	0.0616	0.01572
3	12	0.0484	0.16	0.0406
7	12	0.00862	0.034	0.00874
11	13	0.02225	0.0731	0.01876
12	14	0.0215	0.0707	0.01816

13	15	0.0744	0.2444	0.06268
14	15	0.0595	0.195	0.0502
12	16	0.0212	0.0834	0.0214
15	17	0.0132	0.0437	0.0444
16	17	0.0454	0.1801	0.0466
17	18	0.0123	0.0505	0.01298
18	19	0.01119	0.0493	0.01142
19	20	0.0252	0.117	0.0298
15	19	0.012	0.0394	0.0101
20	21	0.0183	0.0849	0.0216
21	22	0.0209	0.097	0.0246
22	23	0.0342	0.159	0.0404
23	24	0.0135	0.0492	0.0498
23	25	0.0156	0.08	0.0864
26	25	0	0.0382	0
25	27	0.0318	0.163	0.1764
27	28	0.01913	0.0855	0.0216
28	29	0.0237	0.0943	0.0238
30	17	0	0.0388	0
8	30	0.00431	0.0504	0.514
26	30	0.00799	0.086	0.908
17	31	0.0474	0.1563	0.0399
29	31	0.0108	0.0331	0.0083
23	32	0.0317	0.1153	0.1173
31	32	0.0298	0.0985	0.0251
27	32	0.0229	0.0755	0.01926
15	33	0.038	0.1244	0.03194
19	34	0.0752	0.247	0.0632
35	36	0.00224	0.0102	0.00268
35	37	0.011	0.0497	0.01318
33	37	0.0415	0.142	0.0366
34	36	0.00871	0.0268	0.00568

34	37	0.00256	0.0094	0.00984
38	37	0	0.0375	0
37	39	0.0321	0.106	0.027
37	40	0.0593	0.168	0.042
30	38	0.00464	0.054	0.422
39	40	0.0184	0.0605	0.01552
40	41	0.0145	0.0487	0.01222
40	42	0.0555	0.183	0.0466
41	42	0.041	0.135	0.0344
43	44	0.0608	0.2454	0.06068
34	43	0.0413	0.1681	0.04226
44	45	0.0224	0.0901	0.0224
45	46	0.04	0.1356	0.0332
46	47	0.038	0.127	0.0316
46	48	0.0601	0.189	0.0472
47	49	0.0191	0.0625	0.01604
42	49	0.0715	0.323	0.086
42	49	0.0715	0.323	0.086
45	49	0.0684	0.186	0.0444
48	49	0.0179	0.0505	0.01258
49	50	0.0267	0.0752	0.01874
49	51	0.0486	0.137	0.0342
51	52	0.0203	0.0588	0.01396
52	53	0.0405	0.1635	0.04058
53	54	0.0263	0.122	0.031
49	54	0.073	0.289	0.0738
49	54	0.0869	0.291	0.073
54	55	0.0169	0.0707	0.0202
54	56	0.00275	0.00955	0.00732
55	56	0.00488	0.0151	0.00374
56	57	0.0343	0.0966	0.0242
50	57	0.0474	0.134	0.0332

56	58	0.0343	0.0966	0.0242
51	58	0.0255	0.0719	0.01788
54	59	0.0503	0.2293	0.0598
56	59	0.0825	0.251	0.0569
56	59	0.0803	0.239	0.0536
55	59	0.04739	0.2158	0.05646
59	60	0.0317	0.145	0.0376
59	61	0.0328	0.15	0.0388
60	61	0.00264	0.0135	0.01456
60	62	0.0123	0.0561	0.01468
61	62	0.00824	0.0376	0.0098
63	59	0	0.0386	0
63	64	0.00172	0.02	0.216
64	61	0	0.0268	0
38	65	0.00901	0.0986	1.046
64	65	0.00269	0.0302	0.38
49	66	0.018	0.0919	0.0248
49	66	0.018	0.0919	0.0248
62	66	0.0482	0.218	0.0578
62	67	0.0258	0.117	0.031
65	66	0	0.037	0
66	67	0.0224	0.1015	0.02682
65	68	0.00138	0.016	0.638
47	69	0.0844	0.2778	0.07092
49	69	0.0985	0.324	0.0828
68	69	0	0.037	0
69	70	0.03	0.127	0.122
24	70	0.00221	0.4115	0.10198
70	71	0.00882	0.0355	0.00878
24	72	0.0488	0.196	0.0488
71	72	0.0446	0.18	0.04444
71	73	0.00866	0.0454	0.01178

70	74	0.0401	0.1323	0.03368
70	75	0.0428	0.141	0.036
69	75	0.0405	0.122	0.124
74	75	0.0123	0.0406	0.01034
76	77	0.0444	0.148	0.0368
69	77	0.0309	0.101	0.1038
75	77	0.0601	0.1999	0.04978
77	78	0.00376	0.0124	0.01264
78	79	0.00546	0.0244	0.00648
77	80	0.017	0.0485	0.0472
77	80	0.0294	0.105	0.0228
79	80	0.0156	0.0704	0.0187
68	81	0.00175	0.0202	0.808
81	80	0	0.037	0
77	82	0.0298	0.0853	0.08174
82	83	0.0112	0.03665	0.03796
83	84	0.0625	0.132	0.0258
83	85	0.043	0.148	0.0348
84	85	0.0302	0.0641	0.01234
85	86	0.035	0.123	0.0276
86	87	0.02828	0.2074	0.0445
85	88	0.02	0.102	0.0276
85	89	0.0239	0.173	0.047
88	89	0.0139	0.0712	0.01934
89	90	0.0518	0.188	0.0528
89	90	0.0238	0.0997	0.106
90	91	0.0254	0.0836	0.0214
89	92	0.0099	0.0505	0.0548
89	92	0.0393	0.1581	0.0414
91	92	0.0387	0.1272	0.03268
92	93	0.0258	0.0848	0.0218
92	94	0.0481	0.158	0.0406

93	94	0.0223	0.0732	0.01876
94	95	0.0132	0.0434	0.0111
80	96	0.0356	0.182	0.0494
82	96	0.0162	0.053	0.0544
94	96	0.0269	0.0869	0.023
80	97	0.0183	0.0934	0.0254
80	98	0.0238	0.108	0.0286
80	99	0.0454	0.206	0.0546
92	100	0.0648	0.295	0.0472
94	100	0.0178	0.058	0.0604
95	96	0.0171	0.0547	0.01474
96	97	0.0173	0.0885	0.024
98	100	0.0397	0.179	0.0476
99	100	0.018	0.0813	0.0216
100	101	0.0277	0.1262	0.0328
92	102	0.0123	0.0559	0.01464
101	102	0.0246	0.112	0.0294
100	103	0.016	0.0525	0.0536
100	104	0.0451	0.204	0.0541
103	104	0.0466	0.1584	0.0407
103	105	0.0535	0.1625	0.0408
100	106	0.0605	0.229	0.062
104	105	0.00994	0.0378	0.00986
105	106	0.014	0.0547	0.01434
105	107	0.053	0.183	0.0472
105	108	0.0261	0.0703	0.01844
106	107	0.053	0.183	0.0472
108	109	0.0105	0.0288	0.0076
103	110	0.03906	0.1813	0.0461
109	110	0.0278	0.0762	0.0202
110	111	0.022	0.0755	0.02
110	112	0.0247	0.064	0.062

17	113	0.00913	0.0301	0.00768
32	113	0.0615	0.203	0.0518
32	114	0.0135	0.0612	0.01628
27	115	0.0164	0.0741	0.01972
114	115	0.0023	0.0104	0.00276
68	116	0.00034	0.00405	0.164
12	117	0.0329	0.14	0.0358
75	118	0.0145	0.0481	0.01198
76	118	0.0164	0.0544	0.01356

C.4. 44-Bus Samsun

The single line diagram plus the load and line parameters for the 44-Bus practical distribution test system used in Section 7.3.2. of Chapter 7 are presented here.

C.4.1. Load Parameters

The three phase loads for the load buses on the network are presented in Table C.10 below.

Table C.10: Load Parameters

Node	P+jQ [kW,kVar] a	P+jQ [kW,kVar] b	P+jQ [kW,kVar] c
6	0.107+0.06i	0.104+0.058i	0.108+0.061i
8	0.12+0.053i	0.12+0.053i	0.12+0.053i
10	0.214+0.093i	0.214+0.093i	0.214+0.093i
12	0.162+0.084i	0.167+0.089i	0.165+0.089i
13	0.169+0.087i	0.166+0.087i	0.165+0.087i
14	0.238+0.131i	0.246+0.131i	0.241+0.131i
16	0.062+0.024i	0.062+0.024i	0.062+0.024i

18	0.071+0.032i	0.071+0.032i	0.071+0.032i
20	0.327+0.172i	0.327+0.172i	0.327+0.172i
22	0.241+0.133i	0.241+0.133i	0.241+0.133i
24	0.064+0.037i	0.064+0.037i	0
26	0.131+0.07i	0.131+0.07i	0.131+0.07i
28	0.097+0.052i	0.097+0.052i	0.097+0.052i
30	0.07+0.038i	0	0.07+0.038i
33	0.036+0.017i	0	0
35	0.069+0.038i	0.069+0.038i	0
37	0.0730+0.031i	0.0730+0.031i	0.0730+0.031i
39	0	0.067+0.034i	0.065+0.033i
41	0.139+0.072i	0.139+0.072i	0.139+0.072i
43	0.02+0.01i	0	0

C.4.2. Line Parameters

The line parameters and the transformer configurations for the network are presented below.

Table C.11: Line Parameters

Sending Node	Receiving Node	Receiving end voltage	Length (m)	Cable Configuration	Transformer configuration
1	2	34.5	0	100	Y-Y
2	3	34.5	1000	113	0

3	4	34.5	1000	114	0
4	5	34.5	263	135	0
5	6	0.4	0	104	Δ-Y
4	7	34.5	310	195	0
7	8	0.4	0	104	Δ-Y
4	9	34.5	741	195	0
9	10	0.4	0	105	Δ-Y
9	11	34.5	670	195	0
11	12	0.4	0	105	Δ-Y
11	13	0.4	0	105	Δ-Y
11	14	0.4	0	106	Δ-Y
9	15	34.5	354	195	0
15	16	0.4	0	102	Δ-Y
15	17	34.5	280	195	0
17	18	0.4	0	102	Δ-Y
17	19	34.5	279	150	0
17	21	34.5	475	195	0
21	22	0.4	0	105	Δ-Y
21	23	34.5	643	125	0
23	24	0.4	0	102	Δ-Y
23	25	34.5	562	150	0
25	26	0.4	0	104	Δ-Y

23	27	34.5	268	125	0
27	28	0.4	0	103	Δ-Y
27	29	34.5	461	135	0
29	30	0.4	0	101	Δ-Y
2	31	34.5	1000	112	0
31	32	34.5	365	135	0
32	33	0.4	0	104	Δ-Y
31	34	34.5	387	150	0
34	35	0.4	0	104	Δ-Y
34	36	34.5	195	150	0
36	37	0.4	0	104	Δ-Y
31	38	34.5	240	150	0
38	39	0.4	0	104	Δ-Y
38	40	34.5	506	150	0
40	41	0.4	0	106	Δ-Y
40	42	34.5	524	125	0
42	43	0.4	0	101	Δ-Y
2	44	34.5	1000	113	0

The configuration which shows the impedance values for each connection are presented below.

PhaseA	**PhaseB**	**PhaseC**
100=		
0.01+0.08i	0+0.0i	0+0.0i
0+0.0i	0.01+0.08i	0+0.0i
0+0.0i	0+0.0i	0.01+0.08i
101=		
0.044+0.08i	0+0.0i	0+0.0i
0+0.0i	0.044+0.08i	0+0.0i
0+0.0i	0+0.0i	0.044+0.08i
102=		
0.022+0.04i	0+0.0i	0+0.0i
0+0.0i	0.022+0.04i	0+0.0i
0+0.0i	0+0.0i	0.022+0.04i
103=		
0.01746032+0.031746i	0+0.0i	0+0.0i
0+0.0i	0.01746032+0.0317460i	0+0.0i
0+0.0i	0+0.0i	0.01746032+0.031746i
104=		
0.01375+0.025i	0+0.0i	0+0.0i
0+0.0i	0.01375+0.025i	0+0.0i
0+0.0i	0+0.0i	0.01375+0.025i
105=		
0.0088+0.016i	0+0.0i	0+0.0i
0+0.0i	0.0088+0.016i	0+0.0i
0+0.0i	0+0.0i	0.0088+0.016i
106=		
0.006875+0.0125i	0+0.0i	0+0.0i
0+0.0i	0.006875+0.0125i	0+0.0i
0+0.0i	0+0.0i	0.006875+0.0125i
100=		
0.00011+0.0002i	0+0.0i	0+0.0i
0+0.0i	0.00011+0.0002i	0+0.0i
0+0.0i	0+0.0i	0.00011+0.0002i
112=		
0.024523495+0.015298i	0+0.0i	0+0.0i
0+0.0i	0.024523495+0.015298i	0+0.0i
0+0.0i	0+0.0i	0.024523495+0.01529813i
113=		
0.0903342155+0.321084i	0+0.0i	0+0.0i
0+0.0i	0.0903342155+0.321084i	0+0.0i
0+0.0i	0+0.0i	0.0903342155+0.321084i

PhaseA	**PhaseB**	**PhaseC**
114=		
0.14179256+0.16283049i	0+0.0i	0+0.0i
0+0.0i	0.14179256+0.16283049i	0+0.0i
0+0.0i	0+0.0i	0.14179256+0.16283049i
125=		
1.96+4.599919963i	0.59+1.344i	0.59+1.344i
0.59+1.344i	1.96+4.599919963i	0.59+1.344i
0.59+1.344i	0.59+1.344i	1.96+4.599919963i
135=		
0.524+0.1162389282i	0.223+0.0567i	0.223+0.0567i
0.223+0.0567i	0.524+0.1162389282i	0.223+0.0567i
0.223+0.0567i	0.223+0.0567i	0.524+0.1162389282i
150=		
0.494+0.09770353153i	0.223+0.0567i	0.223+0.0567i
0.223+0.0567i	0.494+0.09770353153i	0.223+0.0567i
0.223+0.0567i	0.223+0.0567i	0.494+0.09770353153i
195=		
0.247+0.093619461i	0.123+0.0567i	0.123+0.0567i
0.123+0.0567i	0.247+0.093619461i	0.123+0.0567i
0.123+0.0567i	0.123+0.0567i	0.247+0.093619461i

C.4.3. Single Line Diagram of the 44-bus System

The single line diagram of the 44-bus Samsun network is presented in Fig. C.4 below.

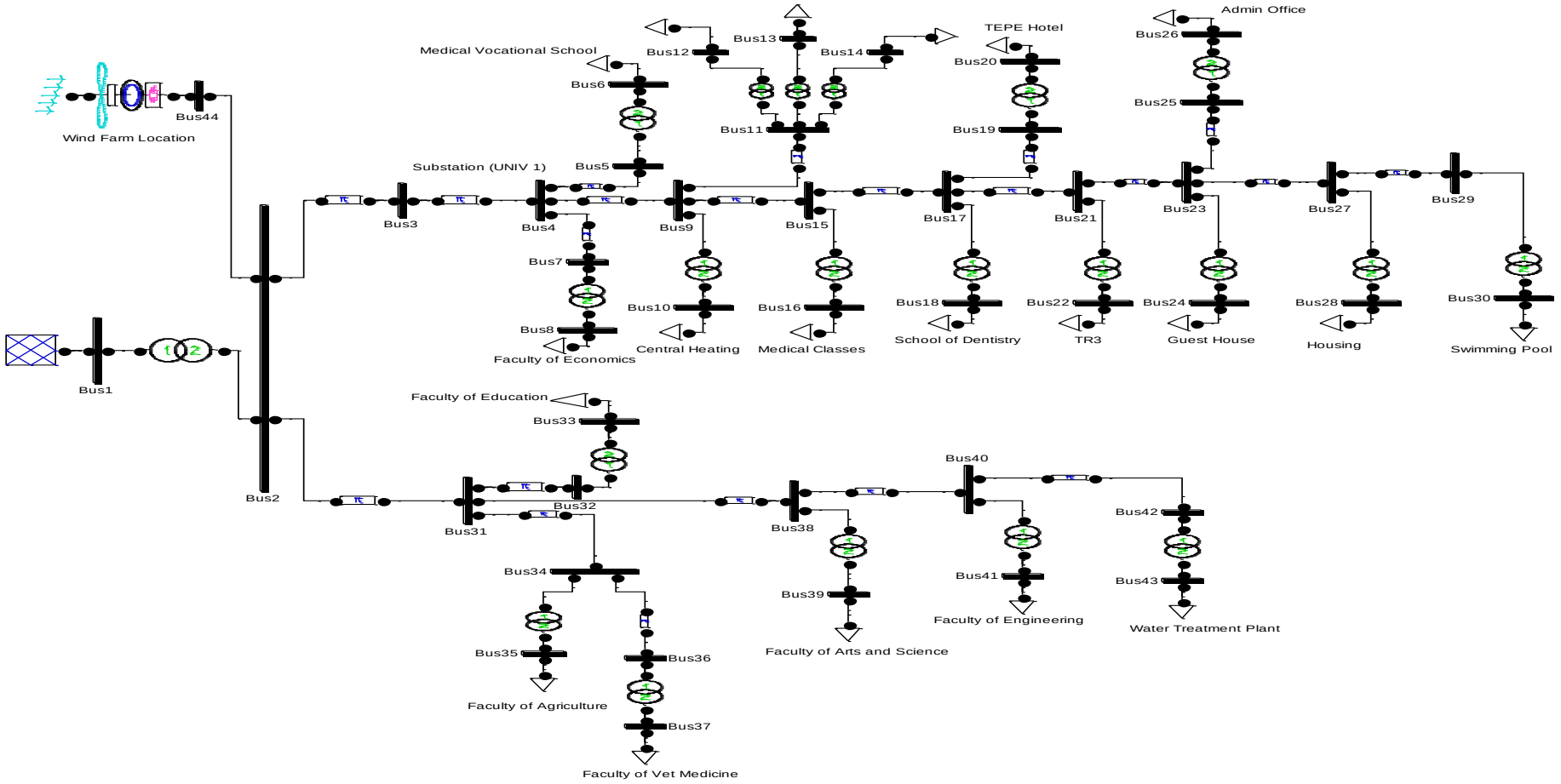

Fig C.4: Single Line Diagram of Samsun Test System

C.5. References

[C.1] Power Systems Test Case Archive, [Online]. Available http://www.ee.washington.edu/research/pstca/

[C.2] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER's Extensible Optimal Power Flow Architecture," *Power and Energy Society General Meeting, 2009 IEEE*, pp. 1-7, July 26-30 2009.